



ESCUELA SUPERIOR POLITECNICA DE CHIMBORAZO

FACULTAD DE CIENCIAS

ESCUELA DE FÍSICA Y MATEMÁTICA

INGENIERÍA ESTADÍSTICA INFORMÁTICA

TESIS DE GRADO

**“ELABORACIÓN DE UN MODELO ESTADÍSTICO QUE AYUDE A
PRONOSTICAR EL VOLUMEN TOTAL DE APROVECHAMIENTO DE LA
MADERA DE PINO EN EL MINISTERIO DEL AMBIENTE CHIMBORAZO”**

**PREVIA A LA OBTENCIÓN DEL TÍTULO DE
INGENIERO EN ESTADÍSTICA INFORMÁTICA**

PRESENTADO POR:

CARLOS MIRANDA CHINLLI

RIOBAMBA – ECUADOR

2013

NOMBRE

FIRMA

FECHA

**Dr. Silvio Álvarez
DECANO**

**Dra. Jenny Orbe
DIRECTORA ESCUELA
FÍSICA Y MATEMÁTICA**

**Dr. Jorge Congacha
DIRECTOR DE TESIS**

**Mat. Marcelo Cortéz
MIEMBRO DEL TRIBUNAL**

**Lcdo. Carlos Rodríguez
DIRECTOR CENTRO
DE DOCUMENTACIÓN**

Yo, CARLOS MIRANDA CHINLLI, soy responsable de las ideas, doctrinas y resultados expuestos en esta Tesis, y el patrimonio intelectual de la Tesis de Grado pertenece a la Escuela Superior Politécnica de Chimborazo.

CARLOS MIRANDA CHINLLI

AGRADECIMIENTO

A ti, Dios todo Poderoso te doy gracias
por darme vida, sabiduría y fuerza, a mis
padres por su apoyo incondicional.

DEDICATORIA

Con mucho cariño a mi madre Zoila Chinlli y a la vida.

ÍNDICE GENERAL

AGRADECIMIENTO.....	iv
DEDICATORIA.....	v
ÍNDICE ABREVIATURAS	xi
ÍNDICE DE TABLAS	xii
ÍNDICE DE GRÁFICOS	xiii
ÍNDICE DE ANEXO.....	xv
INTRODUCCIÓN	xvi
ANTECEDENTES	xvii
JUSTIFICACIÓN	xviii
OBJETIVOS	xviii
CAPITULO I	- 20 -
1. FUNDAMENTO TEÓRICO	- 20 -
1.1.1 POBLACIÓN	- 20 -
1.1.2 MUESTRA.....	- 20 -
1.1.3 DATOS ATÍPICOS (OUTLIERS).....	- 20 -
1.1.4 APROVECHAMIENTO MADERERO.....	- 21 -
1.1.5 MUESTREO FORESTAL.....	- 21 -
1.1.6 TIPOS DE MUESTREO	- 22 -
1.1.6.1 MUESTREO AL AZAR SIMPLE.....	- 22 -
1.1.6.2 MUESTREO ESTRATIFICADO	- 23 -
1.2 VARIABLES DE RODAL (SITIO FORESTAL).....	- 23 -
1.2.1 PINO (PINUS RADIATA)	- 23 -
1.2.3 DIÁMETRO A LA ALTURA DEL PECHO (DAP)	- 24 -
1.2.3 ALTURA (H)	- 25 -
1.2.4 FACTOR DE FORMA (FF).....	- 25 -
1.2.5 VOLUMEN (V).....	- 26 -
1.3 TIPOS DE DATOS	- 26 -
1.3.1 SERIES DE TIEMPO	- 27 -
1.3.1.1 COMPONENTES DE LAS SERIES DE TIEMPO	- 27 -
1.3.1.2 TENDENCIA.....	- 27 -
1.3.1.3 ESTACIONALIDAD	- 28 -

1.3.1.4 CÍCLICIDAD	- 29 -
1.3.1.4 ALEATORIEDAD.....	- 30 -
1.3.2 DATOS DE EJE TRANSVERSAL.....	- 31 -
1.3.3 DATOS COMBINADOS	- 31 -
1.4 MODELOS ESTADÍSTICOS	- 32 -
1.5 PRONÓSTICOS.....	- 32 -
1.5.1 LARGO O CORTO PLAZO	- 33 -
1.5.2 MICRO Y MACRO.....	- 33 -
1.5.3 CUANTITATIVOS O CUALITATIVOS.....	- 33 -
1.5.4 PASOS PARA LOS PRONÓSTICOS	- 34 -
1.6 ERROR DEL PRONÓSTICO	- 36 -
1.6.1 MEDIDAS DE PRECISIÓN	- 37 -
1.6.1.1 LA DESVIACIÓN ABSOLUTA MEDIA (MAD,.....	- 37 -
1.6.1.2 ERROR CUADRÁTICO MEDIO.....	- 37 -
1.6.1.3 ERROR PORCENTUAL ABSOLUTO MEDIO	- 38 -
1.6.1.4 ERROR PORCENTUAL MEDIO.....	- 38 -
1.7 ANÁLISIS EXPLORATORIO DE DATOS (A.E.D)	- 40 -
1.7.1 MEDIDAS DE TENDENCIA CENTRAL	- 40 -
1.7.1.1 MEDIA	- 40 -
1.7.1.2 MEDIANA	- 41 -
1.7.1.3 MODA.....	- 41 -
1.7.2 MEDIDAS DE VARIABILIDAD O DISPERSIÓN.....	- 41 -
1.7.2.1 VARIANZA MUESTRAL.....	- 41 -
1.7.2.2 COEFICIENTE DE VARIACIÓN	- 42 -
1.7.3 MEDIDA DE FORMA	- 42 -
1.7.3.1 COEFICIENTE DE ASIMETRÍA.....	- 42 -
1.7.3.2 COEFICIENTE DE CURTOSIS	- 43 -
1.8 INFERENCIA ESTADÍSTICA.....	- 44 -
1.8.1 ESTIMACIÓN DE PARÁMETROS.....	- 45 -
1.8.1.1 ESTIMACIÓN PUNTUAL	- 45 -

1.8.1.2 ESTIMACIÓN POR INTERVALO.....	- 45 -
1.8.2 PRUEBA DE HIPÓTESIS	- 47 -
1.8.2 .1 VALOR P.....	- 49 -
1.8.3 MÍNIMOS CUADRADOS ORDINARIOS (MCO).....	- 49 -
1.8.4 PROPIEDADES DE LOS ESTIMADORES DE MÍNIMOS CUADRADOS.....	- 52 -
1.8.4.1 ESTIMADOR INSESGADO	- 52 -
1.8.4.2 EFICIENCIA	- 53 -
1.8.4.3 ERROR CUADRÁTICO MEDIO MÍNIMO	- 53 -
1.8.4.4 CONSISTENCIA	- 54 -
1.9 REGRESIÓN MÚLTIPLE	- 54 -
1.9.1 SIGNIFICANCIA DE LA REGRESIÓN.....	- 55 -
1.9.2 PRUEBA F PARA LA SIGNIFICANCIA DE LA REGRESIÓN	- 56 -
1.9.3 PRUEBAS INDIVIDUALES DEL MODELO BÁSICO DE REGRESIÓN MÚLTIPLE. .	- 58 -
1.9.4 COEFICIENTE DE CORRELACIÓN.....	- 60 -
1.9.5 COEFICIENTE DE DETERMINACIÓN.....	- 62 -
1.9.6 COEFICIENTE DE DETERMINACIÓN AJUSTADO \bar{R}^2	- 63 -
1.10 EXTENSIÓN DEL MODELO DE REGRESIÓN LINEAL MÚLTIPLE:.....	- 63 -
1.10.1 MULTICOLINEALIDAD	- 63 -
1.10.2 DETECCIÓN DE LA MULTICOLINEALIDAD	- 64 -
1.10.2.1 ALGUNAS POSIBLES CORRECCIONES A LA MULTICOLINEALIDAD	- 64 -
1.10.2.2 ELIMINAR VARIABLES	- 65 -
1.10.2.3 AUMENTAR EL TAMAÑO MUESTRAL	- 65 -
1.10. 2.4 TRANSFORMACIÓN DE LOS DATOS	- 66 -
1.10.3 HETEROCEDASTICIDAD	- 66 -
1.10.4 DETECCIÓN DE LA HETEROCEDASTICIDAD.....	- 67 -
1.10.4.1 MÉTODO GRÁFICO	- 67 -
1.10.4.2 LA PRUEBA DE WHITE	- 67 -
1.11. ANALISIS DE RESIDUALES	- 68 -

1.12 ELIMINACIÓN PROGRESIVA	- 69 -
1.13 ESPACIO TRIDIMENSIONAL.....	- 69 -
1.13.1 EJES COORDENADOS.....	- 70 -
CAPITULO II	- 71 -
2. MARCO METODOLÓGICO	- 71 -
2.1 PROCEDIMIENTOS PARA ANÁLISIS DE DATOS Y ELABORACIÓN DEL MODELO-	71
-	
2.3 APLICACIÓN DE A.E.D.	- 71 -
2.4 SELECCIÓN DE LAS VARIABLES INCLUIDAS EN UNA REGRESIÓN	- 72 -
MÚLTIPLE.....	- 72 -
2.5 PRONÓSTICOS.....	- 73 -
2.6. HIPÓTESIS ESTADÍSTICO	- 73 -
CAPÍTULO III	- 75 -
3. ANÁLISIS DE RESULTADOS.....	- 75 -
3.1 TEST DE NORMALIDAD	- 75 -
3.2 TRANSFORMACIÓN DE DATOS.....	- 79 -
3.3 ANÁLISIS DESCRIPTIVO DE LAS VARIABLES SIGNIFICATIVAS	- 87 -
3.4 ANÁLISIS DESCRIPTIVO DE LOS DATOS ATÍPICOS	- 90 -
3.5 ANÁLISIS DE REGRESIÓN	- 93 -
3.5.1 ESTUDIO DE LA MATRIZ DE CORRELACIÓN	- 93 -
3.5.2 RESULTADOS REGRESIÓN MÚLTIPLE.....	- 94 -
3.5.3 MODELO BÁSICO DE REGRESIÓN LINEAL.....	- 97 -
3.5.4 ANÁLISIS DE RESIDUOS	- 98 -
3.5.5 PRUEBA DE WHITE	- 100 -
3.6 PRONÓSTICOS Y EVALUACIÓN CON MEDIDAS DE PRECISIÓN	- 101 -
3.6.1 PRONÓSTICO A LARGO PLAZO	- 101 -
3.6.2 PRONÓSTICO MEDIO PLAZO	- 102 -
3.6.3 PRONÓSTICO A CORTO PLAZO.....	- 103 -
3.7. PRUEBA DE HIPÓTESIS PARA LA VARIABLE VOLUMEN V	- 104 -

CAPÍTULO V	- 107 -
4.1 CONCLUSIONES.....	- 107 -
4.2 RECOMENDACIONES	- 109 -
RESUMEN	- 110 -
ABSTRACT	- 111 -
BIBLIOGRAFÍA	- 113 -
ANEXO 1. MATRIZ DE DATOS MENSUALES	- 118 -

ÍNDICE ABREVIATURAS

AED	Análisis exploratorio de datos
DAP	Diámetro a la altura del pecho
DNF	Dirección nacional forestal
FF	Factor de forma
H	Altura
Ln DAP	Logaritmo natural del diámetro a la altura del pecho
Ln FF	Logaritmo natural de factor de forma
Ln NA	Logaritmo natural del número de árboles
Ln H	Logaritmo natural de la altura
Ln V	Logaritmo natural del volumen
MAD	Desviación absoluta media
MAE	Ministerio del ambiente del Ecuador
MAPE	Error porcentual absoluto medio
MBRL	Modelo básico de regresión lineal
MCO	Mínimos cuadrados ordinarios
MELI	Mejor estimador lineal insesgado
MES	Cuadrado medio de residuales
MPE	Error porcentual medio
MSE	Error cuadrático medio
MSR	Cuadrado medio de la regresión
N A	Número de arboles
SSE	Suma de cuadrados de los residuales
SSR	Suma de cuadrados de la regresión
SST	Suma de cuadrados totales
V	Volumen en m^3

ÍNDICE DE TABLAS

TABLA 1. RESULTADO DE UNA PRUEBA DE HIPÓTESIS	- 49 -
TABLA 2. ANOVA PARA REGRESIÓN MÚLTIPLE	- 56 -
TABLA 3. INTERPRETACIÓN DE LOS CONTRASTES.....	- 60 -
TABLA 4. ESTADÍSTICAS DESCRIPTIVAS PARA VOLUMEN V	- 88 -
TABLA 5. ESTADÍSTICAS DESCRIPTIVAS PARA DIÁMETRO A LA ALTURA DEL PECHO DAP ..	- 89 -
TABLA 6. ESTADÍSTICAS DESCRIPTIVAS PARA NÚMERO DE ÁRBOLES NA	- 90 -
TABLA 7. ESTADÍSTICAS DESCRIPTIVAS PARA VOLUMEN V	- 91 -
TABLA 8. ESTADÍSTICAS DESCRIPTIVAS PARA DIÁMETRO A LA ALTURA DEL PECHO DAP ..	- 92 -
TABLA 9. ESTADÍSTICAS DESCRIPTIVAS PARA NÚMERO DE ÁRBOLES NA	- 92 -
TABLA 10. MATRIZ DE CORRELACIÓN.....	- 93 -
TABLA 11. RESULTADOS DE LA REGRESIÓN MÚLTIPLE	- 95 -
TABLA 12. RESULTADOS DE LA REGRESIÓN MÚLTIPLE	- 96 -
TABLA 13. RESULTADOS DE LA REGRESIÓN MÚLTIPLE PARA ERRORES.....	- 100 -
TABLA 14. RESULTADO DE PRONÓSTICO A LARGO PLAZO	- 101 -
TABLA 15. RESULTADO DE PRONÓSTICO A MEDIO PLAZO	- 102 -
TABLA 16. RESULTADO DE PRONÓSTICO A CORTO PLAZO	- 103 -

ÍNDICE DE GRÁFICOS

FIGURA 1. GRÁFICA DE UNA SERIE DE TIEMPO.....	- 28 -
FIGURA 2. GRÁFICA DE UNA SERIE DE TIEMPO.....	- 29 -
FIGURA 3. GRÁFICA DE UNA SERIE DE TIEMPO.....	- 30 -
FIGURA 4. GRÁFICA DE UNA SERIE DE TIEMPO.....	- 31 -
FIGURA 5. GRÁFICAS DEL COEFICIENTE DE ASIMETRÍA	- 43 -
FIGURA 6. GRÁFICAS DEL COEFICIENTE DE CURTOSIS.....	- 44 -
FIGURA 7. AJUSTES Y RESIDUOS	- 51 -
FIGURA 8. GRÁFICA DE DISTRIBUCIÓN F	- 57 -
FIGURA 9. GRÁFICA DE COEFICIENTE DE CORRELACIÓN	- 61 -
FIGURA 10. ESPACIO TRIDIMENSIONAL	- 70 -
FIGURA 11. GRÁFICA NORMAL DE PROBABILIDAD.....	- 75 -
FIGURA 12. GRÁFICA NORMAL DE PROBABILIDAD.....	- 76 -
FIGURA 13. GRÁFICA NORMAL DE PROBABILIDAD.....	- 77 -
FIGURA 14. GRÁFICA NORMAL DE PROBABILIDAD.....	- 77 -
FIGURA 15. GRÁFICA NORMAL DE PROBABILIDAD.....	- 78 -
FIGURA 16. GRÁFICA NORMAL DE PROBABILIDAD.....	- 80 -
FIGURA 17. GRÁFICA NORMAL DE PROBABILIDAD.....	- 81 -
FIGURA 18. GRÁFICA NORMAL DE PROBABILIDAD.....	- 82 -
FIGURA 19. DIAGRAMA DE CAJA.....	- 83 -
FIGURA 20. DIAGRAMA DE CAJA.....	- 84 -

FIGURA 21. DIAGRAMA DE CAJA.....	- 84 -
FIGURA 22. DIAGRAMA DE CAJA.....	- 85 -
FIGURA 23. DIAGRAMA DE CAJA.....	- 86 -
FIGURA 24. DIAGRAMA DE CAJA.....	- 87 -
FIGURA 25. DIAGRAMA DE DISPERSIÓN	- 94 -
FIGURA 26. DIAGRAMA DE DISPERSIÓN EN 3D	- 98 -
FIGURA 27. GRÁFICA DE RESIDUALES	- 99 -
FIGURA 28. GRÁFICA DE LAS REGIONES, ACEPTACIÓN Y RECHAZO	- 105 -

ÍNDICE DE ANEXO

ANEXO 1. MATRIZ DE DATOS MENSUALES	- 118 -
---	----------------

INTRODUCCIÓN

El Ministerio de Ambiente a través de la dirección nacional forestal (DNF) es la institución encargada de diseñar las políticas ambientales y coordinar las estrategias, los proyectos y programas para el cuidado de los ecosistemas y el aprovechamiento sostenible de los recursos naturales. Propone y define las normas para conseguir la calidad ambiental adecuada, con un desarrollo basado en la conservación y el uso apropiado de la biodiversidad y de los recursos con los que cuenta nuestro país. La gestión ambiental es una responsabilidad de todos, porque la calidad de vida depende de las condiciones ambientales en las que nos desarrollamos. Por este motivo, el Ministerio se encarga de recopilar la información de carácter ambiental como un instrumento para educar a la población sobre los recursos naturales, la biodiversidad que posee el país, la manera más adecuada para conservar y utilizar oportunamente estas riquezas.

Con la finalidad de predecir el aprovechamiento del volumen en m^3 de madera, para los meses futuros, se aplicará las técnicas estadísticas tanto la descriptiva, la gráfica de caja y medidas descriptivas, también la inferencial como (estimación de parámetros, prueba de hipótesis) además la regresión múltiple. Es importante pronosticar para planificar y evaluar el recurso forestal (pino) ya que a pesar de ser un recurso renovable tiene un ritmo de crecimiento que puede ser superado por tasa de aprovechamiento de madera.

ANTECEDENTES

En el año 2010 el MAE realizó estudios sobre el aprovechamiento de la madera, la misma que registro 3689180,19 m^3 de madera provenientes de plantaciones forestales, bosques nativos sistemas agroforestales y regeneración natural, las 10 principales especies aprobadas a nivel nacional, tales como: balsa, eucalipto, pino, laurel, pachaco, teca, pigue, sande, pichango y lechero abarcan el 76,79% mientras que las otras especies autorizadas abarcan el 23,21%.

El pino es la tercera especie que más se autorizó para el aprovechamiento a nivel nacional, con 470493,80 m^3 localizados en 11 provincias de la costa y sierra, mediante la aprobación de 442 programas. El 81,93% se autorizó mediante la aprobación de programas de corta de plantaciones de árboles en 10 provincias, haciendo hincapié en Cotopaxi, Chimborazo, Loja, Pichincha, Azuay y Bolívar, y el volumen restante se autorizó mediante programas de corta de árboles plantados en 7 provincias principalmente en Bolívar, Chimborazo, Tungurahua y Pichincha.

JUSTIFICACIÓN

El presente proyecto de investigación se considera de gran importancia en la medida que involucra, al Ministerio del Ambiente como la autoridad ambiental del país y tiene la responsabilidad de asegurar el manejo sustentable de los recursos naturales, conservación de la biodiversidad, aprovechamiento forestal de madera y control de los bosques nativos.

A través de las técnicas estadísticas en particular la estimación del modelo de regresión múltiple se podrá dar a conocer el pronóstico del volumen de madera de pino, en la Provincia de Chimborazo. Al realizar el modelo estadístico para (madera de pino) se podrá establecer conclusiones que permitan en la toma de decisiones, que ayude a mejorar la administración forestal.

OBJETIVOS

OBJETIVO GENERAL

Realizar un modelo estadístico que ayude a pronosticar el volumen total de aprovechamiento de la madera de pino.

OBJETIVOS ESPECÍFICOS

- Establecer que variables de estudio aportan al modelo
- Determinar el modelo adecuado para la situación.
- Medir la significancia de dicho modelo.
- Pronosticar el volumen de aprovechamiento de la madera de pino.

CAPITULO I

1. FUNDAMENTO TEÓRICO

1.1.1 POBLACIÓN

Es un conjunto de medidas o el recuento de todos los elementos o individuos que presentan una característica común. El término población se usa para denotar el conjunto de elementos del cual se extrae la muestra.

1.1.2 MUESTRA

Se define como un conjunto de medidas o el recuento de una parte de los elementos pertenecientes a una población. Los elementos se seleccionan aleatoriamente, es decir, todos los elementos que componen la población tienen la misma posibilidad de ser seleccionados. Para que una muestra sea representativa de la población se requiere que las unidades sean seleccionadas al azar, ya sea utilizando el sorteo, las tablas de números aleatorios, la selección sistemática o cualquier otro método que sea el azar.

1.1.3 DATOS ATÍPICOS (OUTLIERS)

Llamaremos datos atípicos aquellas observaciones que parecen haberse generado de forma distinta de los datos. Pueden ser causadas por errores de medición o transcripción, cambios en el instrumento de medición o heterogeneidad intrínseca de los elementos observados.

Los datos atípicos no pueden ser caracterizados categóricamente como benéficos o problemáticos sino que deben ser contemplados en el contexto del análisis y deben ser evaluados por los tipos de información que pueden proporcionar.

1.1.4 APROVECHAMIENTO MADERERO

Comprende la acción de corta, eliminación de árboles maderables en pie según las necesidades, su extracción y transportación hasta el consumidor o centros de elaboración, en el marco de los principios generales del manejo forestal sustentable.

1.1.5 MUESTREO FORESTAL

Es el acto de tomar o seleccionar una parte (muestra) del bosque (población), realizar mediciones sobre ella e inferir o aplicar los resultados a todo el bosque.

1.1.6 TIPOS DE MUESTREO

Tenemos los siguientes tipos: muestreo al azar simple, muestreo estratificado, muestreo por conglomerado y muestreo sistemático, solo definiremos los dos primeros tipos de muestreo.

1.1.6.1 MUESTREO AL AZAR SIMPLE

Es el método básico de donde se originan los demás muestreos. La idea básica es que al escoger una muestra de “n” unidades en que se encuentre dividida una población, cada una de ellas debe tener la misma oportunidad de ser elegida. Es algo parecido a lo que sucede al comprar un boleto de lotería, donde todos los billetes tienen la misma probabilidad de salir con el premio.

La idea fundamental del muestreo simple aleatorio es que cada una de las posibles combinaciones de unidades muestrales tiene la misma posibilidad de ser seleccionada. Esto se asegura al diseñar el muestreo de forma tal que la selección

de una unidad en particular (sitio o parcela de muestreo) no está condicionada a la selección de cualquier otra unidad que haya sido seleccionada previamente.

1.1.6.2 MUESTREO ESTRATIFICADO

El muestreo estratificado es un muestreo que toma ventaja de la información adicional que se tiene de la población. Aquí las unidades de muestreo se agrupan por alguna característica en común o estratos. Regularmente en plantaciones, la agrupación se hace por edad de plantación, densidad de plantación o bien alguna otra variable como sistema, o método de plantación. En este caso, se levanta muestras en cada grupo (o estrato) y posteriormente se combinan las estimaciones por grupo para proporcionar el estimador del parámetro de interés de la población. Antes de aplicar este tipo de muestreo es conveniente enfatizar que sólo en el caso de que la variación entre unidades de muestreo de un mismo grupo (estrato) sea menor que entre unidades que no son del mismo grupo, entonces los estimadores de este tipo de muestreo son más precisos que los de un muestreo al azar.

1.2 VARIABLES DE RODAL (SITIO FORESTAL)

1.2.1 PINO (PINUS RADIATA)

Es nativo del norte de México, Islas Guadalupe y el estado de California en USA. Es una especie que se ha introducido en gran escala en Nueva Zelandia, Chile, Australia, Sud África, y en menor escala en muchos otros países, entre ellos Ecuador.

El pino es una conífera que ha demostrado adaptarse a muchas condiciones ambientales, es decir es una especie “plástica” y rústica. La madera es usada principalmente en la fabricación de pulpa y papeles periódicos, posee buena resistencia, y por ello es usada en madera aserrada para construcciones livianas, tableros de fibra, de partículas, contrachapados y en mueblería.

El pino es la especie más plantada en bosques de mayores superficies en el Ecuador. Es una de las pocas especies que bajo condiciones comunes se adapta en la sierra en alturas sobre los 3200 msnm y en suelos marginales para la agricultura.

1.2.3 DIÁMETRO A LA ALTURA DEL PECHO (DAP)

Medida del diámetro de la circunferencia del tronco de un árbol a la altura de 1,30 metros sobre el nivel de suelo. Para la obtención de esta medida se utiliza la forcípula o la cinta diamétrica. La forcípula es más cómodo para medir arboles hasta de 50 cm de DAP, para arboles más gruesos, se utiliza la cinta diamétrica

comparada con la forcípula que proporciona una lectura más exacta. En muchos casos se utiliza la cinta métrica para medir los DAP, en este caso medimos la circunferencia del árbol.

$$2\pi r = \text{Perímetro}$$

$$\pi 2r = \text{Perímetro}$$

$$\pi d = \text{Perímetro}$$

$$d = \frac{\text{Perímetro}}{\pi}$$

1.2.3 ALTURA (H)

La altura de un árbol se define como la distancia del suelo a la punta o ápice del árbol a lo largo del fuste (tronco del árbol). Los instrumentos utilizados para medir las alturas en los árboles son: Hipsómetros, Silva, Blumeleiss, Haga, etc.

1.2.4 FACTOR DE FORMA (FF)

Es el factor de reducción, la obtención de dicho factor se efectúa tomando el volumen real y dividiéndolo entre el volumen de un cilindro con el diámetro medido a una altura de 1.30 metros en el árbol, cada especie tiene característico factor de forma que también varía durante el tiempo de crecimiento

1.2.5 VOLUMEN (V)

Es la cantidad de madera trozada de un árbol o masa boscosa, que rinde según una unidad de medida (metros cúbicos, pie tablares, pulgadas). La fórmula utilizada es:

$$V = \frac{\pi}{4} DAP^2 * FF * H * NA$$

Donde:

V: Volumen comercial del árbol (m^3)

DAP: Diámetro a la altura del pecho (m)

FF: Factor de forma

H: altura comercial de fuste (m)

NA: Número de árboles

1.3 TIPOS DE DATOS

Hay tres tipos de datos disponibles para el análisis empírico: series de tiempo, series de eje transversal e información combinada (combinación de series de tiempo y transversal).

1.3.1 SERIES DE TIEMPO

Llamamos serie de tiempo a datos estadísticos que se recopilan, observan o registran en intervalos de tiempo regulares (diario, semanal, mensual, trimestral, semestral, anual). Dado que los acontecimientos pasados pueden tener influencia sobre acontecimientos futuros el tiempo es un parámetro importante en los conjuntos de series de tiempo.

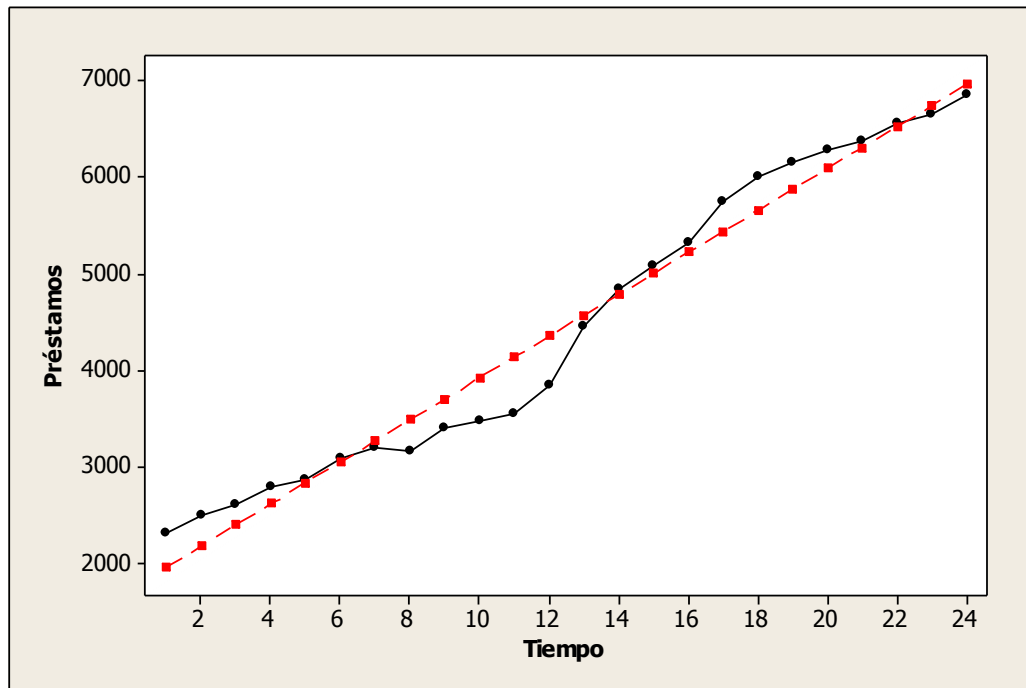
1.3.1.1 COMPONENTES DE LAS SERIES DE TIEMPO

La teoría clásica considera una serie de tiempo formado por cuatro componentes teóricas: tendencia, estacionales, cíclicas y aleatoriedad.

1.3.1.2 TENDENCIA

Es el componente de largo plazo que representa el crecimiento o decrecimiento en la serie de tiempo a lo largo de un periodo. Por ejemplo:

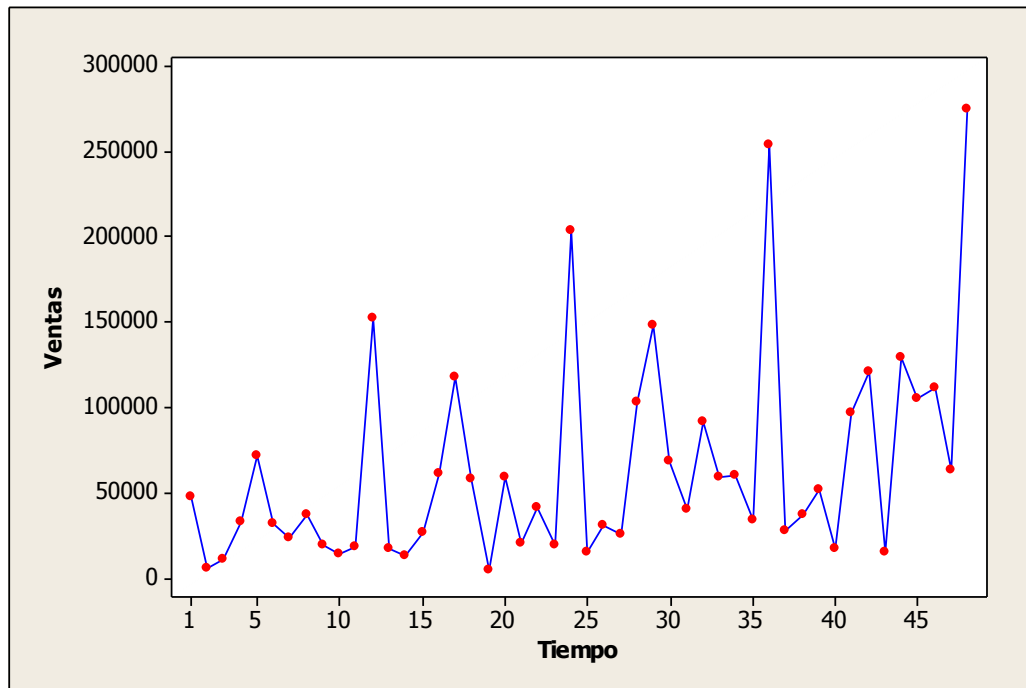
FIGURA 1. GRÁFICA DE UNA SERIE DE TIEMPO



1.3.1.3 ESTACIONALIDAD

Se refiere a un patrón de cambio, dentro de un periodo anual. El componente de la serie de tiempo que representa la variabilidad en los datos debida a la influencia de las estaciones, se llama componente estacional. Esta variación corresponde a los movimientos de la serie que recurren año tras año en los mismos meses (o en los mismos trimestres) del año poco más o menos con la misma intensidad. Por ejemplo:

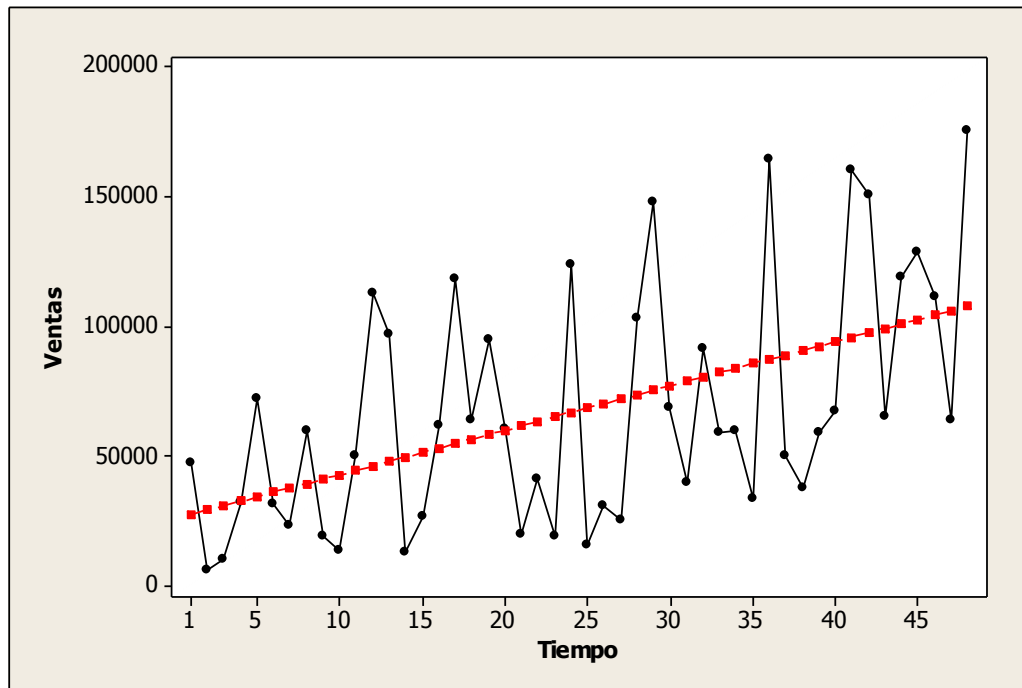
FIGURA 2. GRÁFICA DE UNA SERIE DE TIEMPO



1.3.1.4 CÍCLICIDAD

Son caracterizados por los movimientos oscilatorios alrededor de la tendencia y aproximadamente regulares de los datos, en el largo plazo (periodos mayores que un año), en torno de la tendencia. Los ciclos pueden ser debidos a muchos factores, principalmente a los fenómenos naturales, como estaciones del año, periodos de lluvia o de calor intenso. También pueden ser debidos a fenómenos socioeconómicos, como los periodos de recesión o euforia típicos de los ciclos económicos. Por ejemplo:

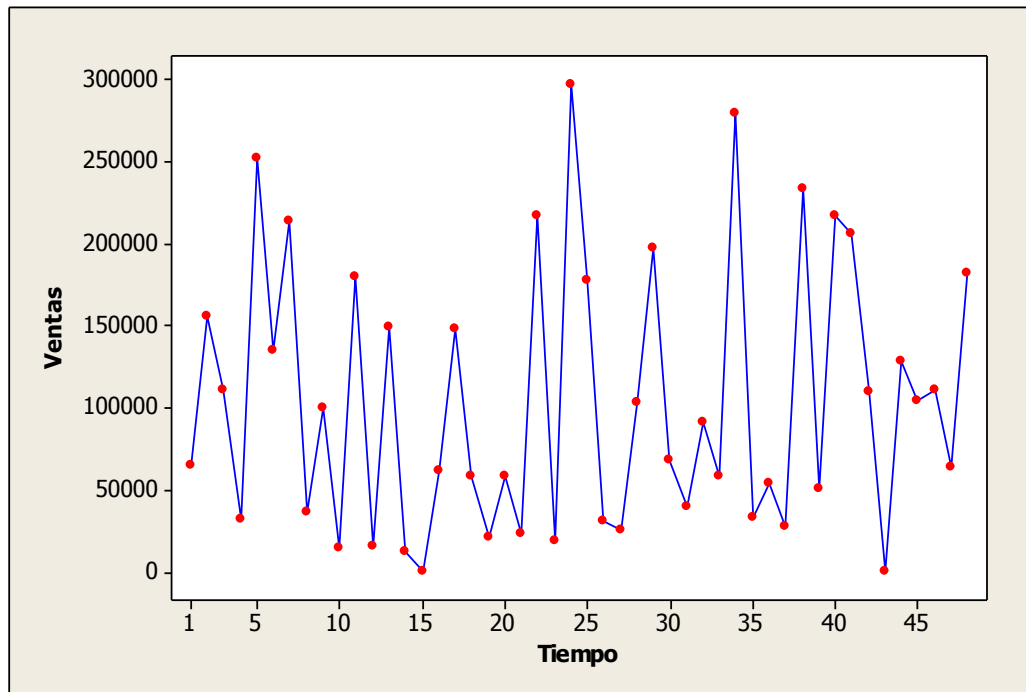
FIGURA 3. GRÁFICA DE UNA SERIE DE TIEMPO



1.3.1.4 ALEATORIEDAD

Las variaciones, llamadas aleatorias pueden ser debidas a causas naturales como sequias, inundaciones, heladas, epidemias y terremotos o causas sociales como periodos de relaciones internacionales y hasta por planes económicos así como pueden ser debidas a eventos no identificados. El efecto provocado por esta componente, en general no mensurable, es de corta duración e intensidad variable. Por ejemplo:

FIGURA 4. GRÁFICA DE UNA SERIE DE TIEMPO



1.3.2 DATOS DE EJE TRANSVERSAL

Los datos transversales consisten en datos de una o más variables recopiladas en el mismo punto del tiempo o en el mismo período. De manera característica, los parámetros de corte transversal puede cambiar con el tiempo de una manera que no se refleja en la elección de variables independientes de series de tiempo o los individuos pueden variar en forma importante dentro del corte trasversal de tal forma que no se refleja en la elección de variables de corte transversal.

1.3.3 DATOS COMBINADOS

Los datos de panel o combinados reúnen elementos de series de tiempo y eje transversal.

1.4 MODELOS ESTADÍSTICOS

Es la abstracción de fenómeno real o proceso que aísla los aspectos y relaciones relevantes, proporcionando una estructura lógica para el sistema de interés, cuyos parámetros son desconocidos, y se estiman a partir de datos empíricos.

1.5 PRONÓSTICOS

Es el proceso de estimación en situaciones de incertidumbre, sobre una actividad futura con las condiciones y variaciones registradas hasta la fecha, lo cual permite planear y tomar decisiones, dentro de las organizaciones grandes y pequeñas, la necesidad de tener pronósticos está en todas las líneas funcionales, aprovechamiento de madera, en las áreas de finanzas, marketing, personal, producción. Los procedimientos de pronósticos podrían clasificarse como:

1.5.1 LARGO O CORTO PLAZO

Los pronósticos de largo plazo son necesarios para establecer el curso general de una organización y son el enfoque exclusivo de la alta dirección.

Los pronósticos de corto plazo se usan para diseñar estrategias inmediatas, y los mandos y las gerencias de primera línea los usan para cubrir las necesidades del futuro inmediato.

1.5.2 MICRO Y MACRO

En la medida que involucre pequeños detalles en comparación con una gran escala. Por ejemplo, al director del MAE le podría interesar el pronóstico del número de trabajadores necesarios durante los próximos meses (un micro pronóstico), mientras que el gobierno quiere pronosticar la cantidad total de personas empleadas en el país entero (un macro pronóstico).

1.5.3 CUANTITATIVOS O CUALITATIVOS

Una técnica cualitativa no requiere manipulación abierta de datos. Solamente se utiliza el juicio de quien pronostica, el juicio de esta persona es resultado de la manipulación mental de datos históricos. Las técnicas cuantitativas no necesitan

elementos de juicio son procedimientos mecánicos que producen resultados cuantitativos. Por supuesto algunos procedimientos cuantitativos requieren una manipulación mucho más sofisticada de los datos que otros. Por ejemplo medición de la altura, volumen etc.

1.5.4 PASOS PARA LOS PRONÓSTICOS

El reconocimiento de que las técnicas de pronósticos operan con los datos generados por sucesos históricos lleva a la identificación de los siguientes cinco pasos en el proceso de pronósticos:

1. Formulación del problema y recolección de datos
2. Manipulación y depuración de datos
3. Construcción y evaluación del modelo
4. Aplicación del modelo (el pronóstico real)
5. Evaluación del pronóstico

En el paso 1, **la formulación del problema y la recolección de datos** se tratan como un paso único, debido a que están íntimamente relacionados. El problema determina los datos adecuados. Si se considera una metodología cuantitativa de pronósticos, los datos relevantes deben estar disponibles y ser correctos. A

menudo, la tarea de obtener y recopilar los datos adecuados es desafiante y se consume mucho tiempo. Si no se dispone de los datos adecuados, el problema podría tener que redefinirse o se tendría que emplear una metodología no cuantitativa de pronósticos.

El paso 2, **manipulación y depuración de datos**, con frecuencia es necesario. En el proceso de pronóstico es posible tener demasiados datos o muy pocos. Algunos datos pueden ser irrelevantes al problema. Otros podrían tener valores omitidos que deban estimarse. Algunos datos podrían tener que expresarse en unidades diferentes de las originales. Por lo general, se requiere cierto esfuerzo para obtener los datos de la forma requerida a fin de utilizar determinados procedimientos de pronósticos.

El paso 3, **construcción y evaluación del modelo**, implica adecuar los datos recolectados en un modelo de pronósticos que sea adecuado en términos de minimización del error de pronósticos. Cuanto más simple sea el modelo, resulta mejor para obtener la aceptación del proceso de pronóstico por parte de los administradores que toman las decisiones en la organización. Obviamente, el juicio intuitivo está implicado en este proceso de selección.

El Paso 4, **aplicación del modelo**, consiste en los pronósticos reales del modelo que se generan una vez que se han recolectado y quizás reducido a solo datos adecuados. Frecuentemente, los pronósticos de periodos recientes en los que se conocen los valores históricos reales se utiliza para verificar la precisión del

proceso. Los errores de pronóstico que se observa y se resume de cierta forma aparecen en el paso siguiente (paso 5).

El paso 5, **evaluación del pronóstico**, implica comparar los valores del pronóstico con los valores históricos reales. Una vez que se ha completado el modelo de pronósticos, se realizan los pronósticos para los diversos periodos y se comparan con los valores históricos conocidos. Algunos procedimientos de pronóstico reporta la suma de los valores absolutos de los errores promedio de pronóstico. Otros procedimientos toman en cuenta la suma de los cuadrados del error, lo cual compara con cifras similares de métodos alternos de pronóstico. Algunos procedimientos también localizan y reportan la magnitud de los términos de los errores a lo largo del periodo de pronóstico. Frecuentemente, el examen de los patrones de errores lleva al analista a modificar el procedimiento de pronósticos.

1.6 ERROR DEL PRONÓSTICO

Con el análisis entre datos proyectados o datos reales, Se incluyen últimas observaciones y se vuelve a estimar el modelo con la muestra y generar sus pronósticos.

El error de pronóstico está dado por :

$$e_i = Y_i - \hat{Y}_i$$

Donde:

e_t , error del pronóstico en el periodo t

Y_t , valor real en el periodo t

\hat{Y}_t , valor del pronóstico para el periodo t

1.6.1 MEDIDAS DE PRECISIÓN

1.6.1.1 LA DESVIACIÓN ABSOLUTA MEDIA (MAD, del inglés mean absolute deviation)

Mide la precisión del pronóstico al promediar las magnitudes de los errores de pronóstico (valores absolutos de cada error). MAD es más útil cuando el analista quiere medir el error de pronósticos en las mismas unidades que la serie original.

$$MAD = \frac{1}{n} \sum_{i=1}^n \left| Y_i - \hat{Y}_i \right|$$

1.6.1.2 ERROR CUADRÁTICO MEDIO (MSE, del inglés mean squared error)

Es otro método para evaluar una técnica de pronóstico. Cada error de pronóstico o residual se eleva al cuadrado, luego se suman y se dividen entre el número de observaciones. Este método penaliza los errores grandes de pronóstico debido a que los errores se elevan al cuadrado, lo cual es importante, una técnica que produce errores moderados podría ser preferible a una que por lo general, tiene errores pequeños, pero en ocasiones produce errores grandes.

$$MSE = \frac{1}{n} \sum_{i=1}^n \left(Y_i - \hat{Y}_i \right)^2$$

1.6.1.3 ERROR PORCENTUAL ABSOLUTO MEDIO (MAPE, del inglés mean absolute porcentaje error)

Este método es útil cuando el tamaño o magnitud de la variable del pronóstico es importante para evaluar la precisión del pronóstico. El MAPE proporciona una indicación de cuán grandes son los errores de pronóstico en comparación con los valores reales de la serie.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{\left| Y_i - \hat{Y}_i \right|}{Y_i} * 100$$

1.6.1.4 ERROR PORCENTUAL MEDIO (MPE, del inglés mean porcentaje error)

A veces es necesario determinar si un método de pronóstico tiene sesgo (produce pronósticos más altos o más bajos de manera sistemática). Si el método de pronóstico no tiene sesgo, el MPE producirán un número cercano a cero. Si el resultado es un alto porcentaje negativo, el método sobrestima de forma consistente, y el resultado es un porcentaje alto positivo, el método subestima consistentemente.

$$MPE = \frac{1}{n} \sum_{i=1}^n \frac{(Y_i - \hat{Y}_i)}{Y_i}$$

Parte de la decisión de utilizar una técnica específica de pronóstico implica la determinación de si esta técnica producirá errores de pronósticos que se juzguen ser suficientemente pequeños. Se puede esperar que una buena técnica de pronóstico produzca errores de pronóstico relativamente pequeños de forma persistente.

Las cuatro medidas de precisión de pronóstico que se ha descrito se utilizan para:

- Comparar la precisión de dos o más técnicas diferentes
- Medir la utilidad o confiabilidad de una técnica específica.
- Ayudar a buscar una técnica óptima

1.7 ANÁLISIS EXPLORATORIO DE DATOS (A.E.D)

Es un conjunto de técnicas estadísticas cuya finalidad es conseguir un entendimiento básico de los datos y de las relaciones existentes entre las variables analizadas. Para conseguir este objetivo el AED proporciona métodos sistemáticos sencillos para organizar y preparar los datos, detectar fallos en el diseño y recogida de los mismos, tratamiento y evaluación de datos ausentes (missing), identificación de casos atípicos (outliers) y comprobación de los supuestos subyacentes en la mayor parte de las técnicas multivariantes (normalidad, linealidad, homocedasticidad).

1.7.1 MEDIDAS DE TENDENCIA CENTRAL

1.7.1.1 MEDIA

Es una medida de tendencia central que denota el promedio de un conjunto de datos. Se calcula como:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

1.7.1.2 MEDIANA

Es una medida de tendencia central, de un conjunto de datos ordenados de manera creciente es el valor para el cual, la mitad de éstos es menor que éste valor y la otra mitad mayor.

1.7.1.3 MODA

Es el valor que ocurre con mayor frecuencia en la muestra. La moda puede no existir, e incluso no ser única en caso de existir.

1.7.2 MEDIDAS DE VARIABILIDAD O DISPERSIÓN

1.7.2.1 VARIANZA MUESTRAL

Es la medida de dispersión más importante, denotada con S^2 está dada por

$$S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$$

La desviación estándar de la muestra, denotada con S , es la raíz cuadrada positiva de s^2 es decir,

$$s = \sqrt{s^2}$$

1.7.2.2 COEFICIENTE DE VARIACIÓN

Permite comparar la dispersión entre dos poblaciones distintas e incluso, compara la variación producto de dos variables diferentes (que puede provenir de una misma población) y es la razón entre la desviación típica o estándar y la media aritmética. Se interpreta como porcentaje, por lo que nos permite decidir entre dos muestras, cual es la que presenta mayor dispersión.

$$CV = \frac{S}{x} * 100\%$$

1.7.3 MEDIDA DE FORMA

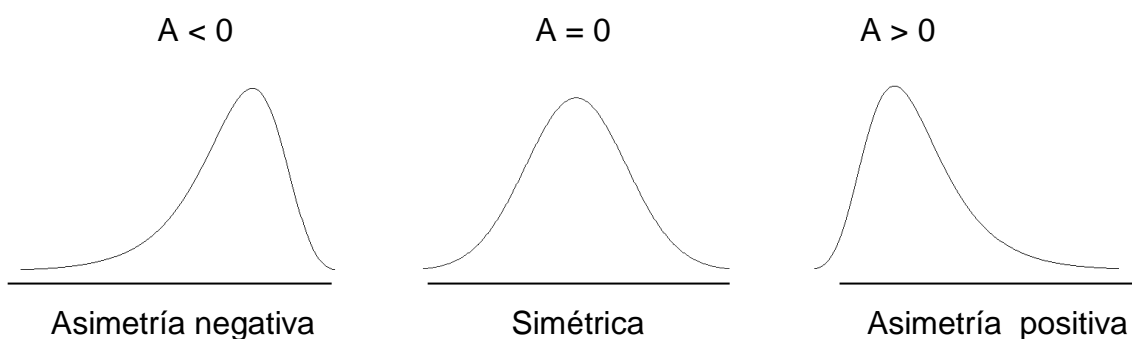
1.7.3.1 COEFICIENTE DE ASIMETRÍA

Es la medida de la forma de la distribución, la simetría es importante para saber si los valores de las variables se concentran en una determinada zona del recorrido de la variable, se calcula por:

$$A = \frac{\sum_{i=1}^n (x_i - \bar{x})^3}{ns^3}$$

Los resultados pueden ser:

FIGURA 5. GRÁFICAS DEL COEFICIENTE DE ASIMETRÍA



1.7.3.2 COEFICIENTE DE CURTOSIS

Es una medida que proporciona qué tan puntiaguda es la distribución de probabilidad, la estadística de curtosis está dado por:

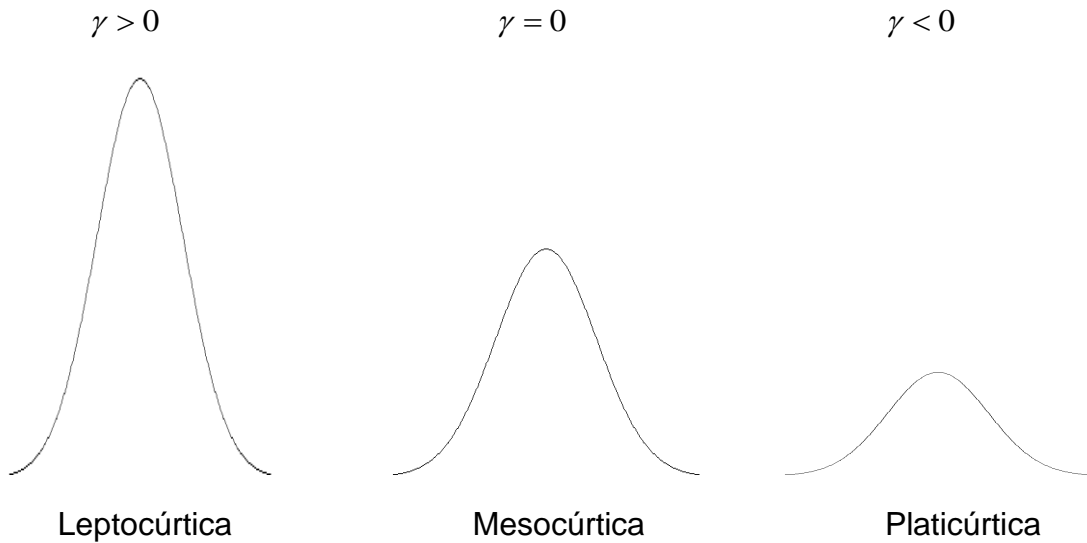
$$\gamma = \frac{\mu_4}{S^4} - 3$$

Donde

$$\mu_4 = \frac{1}{4} \sum_{i=1}^n (X_i - \bar{X})^4$$

Los resultados pueden ser:

FIGURA 6. GRÁFICAS DEL COEFICIENTE DE CURTOSIS



1.8 INFERENCIA ESTADÍSTICA

La teoría de la inferencia estadística consiste de aquellos métodos y procedimientos por los cuales uno hace generalizaciones o conclusiones acerca de una población a partir del estudio de una o más muestras representativas.

1.8.1 ESTIMACIÓN DE PARÁMETROS

Son todas aquellas medidas que describen numéricamente la característica de una población a partir del conocimiento de la muestra. También se denomina valor verdadero, ya que una característica poblacional tendrá un solo parámetro (media, proporción, varianza, etc.). Sin embargo una población puede tener varias características y por tanto, varios parámetros. Hay dos formas de estimar un parámetro: puntual o por intervalos.

1.8.1.1 ESTIMACIÓN PUNTUAL

Un estimador puntual de algún parámetro poblacional θ es un único valor de $\hat{\theta}$. A un estimador también se le conoce con el nombre de estadística. Por ejemplo, \bar{X} , S, etc.

1.8.1.2 ESTIMACIÓN POR INTERVALO

Una estimación por intervalo de un parámetro poblacional θ es un intervalo de la forma $\hat{\theta}_I < \theta < \hat{\theta}_S$, donde $\hat{\theta}_I$ (límite inferior) y $\hat{\theta}_S$ (límite superior) dependen de los

valores del estimador $\hat{\theta}$ de una muestra particular y también de la distribución muestreo de $\hat{\theta}$.

Diferentes muestras, por lo general, darán valores diferentes de $\hat{\theta}$, por lo tanto, valores diferentes de $\hat{\theta}_I$ y $\hat{\theta}_S$, luego los límites del intervalo son valores correspondientes de las variables aleatorias $\hat{\theta}_I$ y $\hat{\theta}_S$. De la distribución muestral de $\hat{\theta}$ seremos capaces de determinar $\hat{\theta}_I$ y $\hat{\theta}_S$ tales que $P(\hat{\theta}_I < \theta < \hat{\theta}_S)$ es igual a cualquier valor positivo fraccionado y especificado cuidadosamente. Si por ejemplo queremos encontrar $\hat{\theta}_I$ y $\hat{\theta}_S$ tales que $P(\hat{\theta}_I < \Theta < \hat{\theta}_S) = 1 - \alpha$

Para $0 < \alpha < 1$, luego tenemos una probabilidad de $1 - \alpha$ de seleccionar una muestra aleatoria que definirá un intervalo que contiene Θ , se llama nivel de significancia y es un valor prefijado. A su valor complemento se lo llama nivel de confianza, es decir $1 - \alpha$. Por ejemplo.

$$P(\bar{X} - Z_{\alpha/2} \frac{S}{\sqrt{n}} < \mu < \bar{X} + Z_{\alpha/2} \frac{S}{\sqrt{n}}) = 1 - \alpha$$

$$\theta_I = \bar{X} - Z_{\alpha/2} \frac{S}{\sqrt{n}}$$

$$\theta_S = \bar{X} + Z_{\alpha/2} \frac{S}{\sqrt{n}}$$

El intervalo $\hat{\theta}_I < \Theta < \hat{\theta}_S$, calculado de la muestra aleatoria seleccionada es entonces llamado intervalo de confianza de $(1 - \alpha)$ 100%, la fracción $1 - \alpha$ se llama coeficiente de confianza o grado de confianza inferior y superior respectivamente.

1.8.2 PRUEBA DE HIPÓTESIS

Una hipótesis estadística es una aseveración o conjetura con respecto a una o más poblaciones.

Existen dos clases de hipótesis estadísticas: nula y alterna. La hipótesis nula se refiere a cualquier hipótesis que queremos probar y se denota por H_0 . El rechazo de H_0 permite la aceptación de una hipótesis denominada hipótesis alterna o alternativa, la misma que se denota por H_1 .

En muchas situaciones estadísticas, que incluye los pronósticos, el interés está en probar cierta afirmación sobre la población más que en estimar o predecir alguno de sus parámetros. A este proceso se llama prueba de hipótesis. Probar las hipótesis implica los siguientes pasos:

1. **Establecimiento de hipótesis**, previamente se decide si una prueba es de una o dos colas la cual o las cuales se definirán por la hipótesis alterna H_1 . Para formular H_1 debemos preguntarnos ¿Qué deseo concluir? o ¿Qué creo que es

verdadero? El complemento de la hipótesis alternativa constituye la hipótesis nula H_0 .

2. **Selección de un nivel de significancia**, el cual debe ser apropiado para la decisión anterior. Generalmente se toman valores del orden 0.01, 0.05 ó 0.10 ó entre 0.01 y 0.10.

3. **Tamaño de la muestra y cálculos de estadístico(s)**, se toma en cuenta el tamaño de la muestra y aspectos de las distribuciones muestrales, como por ejemplo la naturaleza de la población investigada o la forma que se obtuvo ésta. Además se considera el estadístico para decidir la distribución (t o z).

4. **Especificación de las regiones de aceptación y rechazo**, Se esboza a través de un gráfico dichas regiones, delimitadas por el valor o los valores críticos de z o t definidos por el nivel de significancia prefijado.

5. **Decisión estadística**, para tomar una decisión, se compara el valor de la estadística estandarizada calculada con el valor o los valores críticos, o se ve si este valor, cae en la región (es) de aceptación o rechazo (s) y luego dependiendo de esta comparación se acepta o se rechaza la hipótesis nula H_0 .

Cuando se sigue estos pasos, puede ocurrir dos tipos de errores como se indica en la tabla:

TABLA 1. RESULTADO DE UNA PRUEBA DE HIPÓTESIS

Decisión	H_0 Verdadera	H_0 Falsa
Aceptar H_0	Decisión correcta ($1 - \alpha$)	Error tipo II (β)
Rechazar H_0	Error tipo I (α)	Decisión correcta ($1 - \beta$)

1.8.2 .1 VALOR P

El valor p mide la probabilidad de un error tipo I , es decir, la probabilidad de rechazar en forma incorrecta una hipótesis nula verdadera. Entre mayor es el valor de p, es más probable que será un error rechazar la hipótesis nula.

1.8.3 MÍNIMOS CUADRADOS ORDINARIOS (MCO)

Especificado un modelo $Y_i = \beta_1 + \beta_2 X_i + e_i$ si se dispusiese todas las observaciones correspondientes a la población, se podría obtener los valores de los parámetros poblacionales β_1 y β_2 . En la práctica se trabaja con una muestra de tamaño n, extraída de la población. En estas circunstancias, no es posible

obtener los valores de β_1 y β_2 , si no tan solo una aproximación a los mismos $\hat{\beta}_1$, $\hat{\beta}_2$ que dependerán de la muestra concreta con la se esté trabajando. Se tiene por tanto la regresión muestral:

$$\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + \hat{e}_i$$

donde \hat{e}_i , son los residuos, es decir, la diferencia entre el valor observado de la variable que tratamos de explicar y el valor ajustado para la misma.

Sin embargo, para cada muestra seleccionada, los valores correspondientes a β_1 y β_2 en función de las observaciones muestrales, denominaremos a $\hat{\beta}_1$ y $\hat{\beta}_2$ estimadores de los parámetros poblacionales.

Existen varios métodos para la obtención de la expresión matemática de los estimadores, pero el más generalizado es el de los mínimos cuadrados.

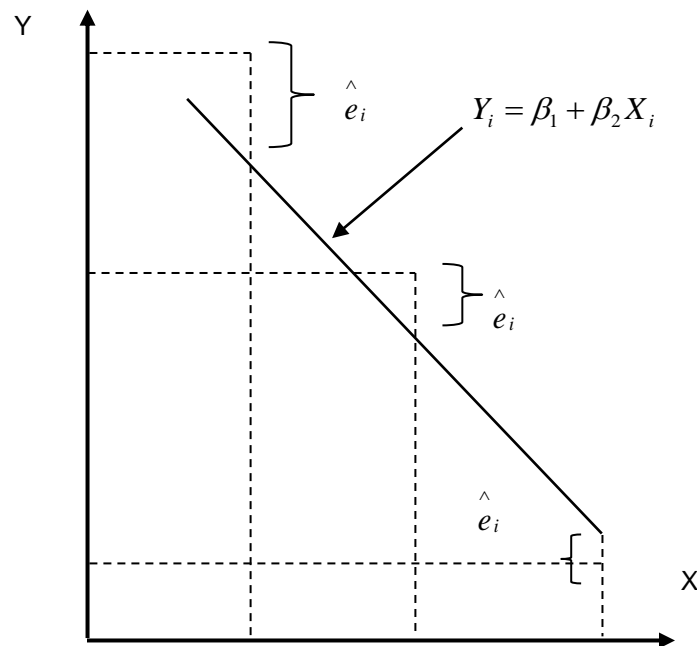
$$\hat{e}_i = Y_i - \hat{Y}_i = Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i$$

Así la suma de los cuadrados de los errores estimados será:

$$\sum \hat{e}_i^2 = \sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)^2$$

La minimización nos permite obtener $\hat{\beta}_1$ y $\hat{\beta}_2$ que son los estimadores de los parámetros poblacionales de β_1 y β_2 . En términos gráficos, el problema queda representado.

FIGURA 7. AJUSTES Y RESIDUOS



Aplicando, por tanto, las condiciones de mínimos se obtienen las siguientes expresiones de los estimadores:

$$\hat{\beta}_2 = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

$$\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{X}$$

donde \bar{X} e \bar{Y} son las medias de las variables independiente y dependiente, respectivamente.

$\hat{\beta}_1$ y $\hat{\beta}_2$ tal como quedan definidos, son los estimadores mínimos cuadrados que, para cada muestra concreta tomara diferentes valores.

1.8.4 PROPIEDADES DE LOS ESTIMADORES DE MÍNIMOS CUADRADOS

En el contexto de la regresión puede probarse que los estimadores de MCO son MELI. Hay cuatro importantes de los estimadores que son:

1.8.4.1 ESTIMADOR INSESGADO

Diríamos que $\hat{\beta}$ es un estimador insesgado si la media o valor esperado de $\hat{\beta}$ es igual al valor verdadero, es decir $E(\hat{\beta}) = \beta$. Para aclarar la exposición definiremos el sesgo asociado con un parámetro estimado como sigue como:

$$Sesgo = E(\hat{\beta}) - \beta$$

Aunque la ausencia de sesgo en un estimador es una propiedad deseable, no supone nada acerca de la dispersión del estimador alrededor del parámetro verdadero. En general, a uno le gustaría que el estimador fuera insesgado y también que tuviera una dispersión muy pequeña alrededor de la media. Esto nos sugiere que debemos definir un segundo criterio que le permitiera elegir entre estimadores insesgados alternativos.

1.8.4.2 EFICIENCIA

Decimos que $\hat{\beta}$ es un estimador insesgado eficiente si para un tamaño muestral dado la varianza de $\hat{\beta}$ es menor que la varianza de cualesquier otros estimadores insesgados.

1.8.4.3 ERROR CUADRÁTICO MEDIO MÍNIMO

En diversas circunstancias nos vemos obligados a sacrificar el sesgo y la varianza de los estimadores. Esto es, cuando la meta de un modelo es maximizar la precisión de las predicciones. Un criterio que es útil a este respecto es la meta de minimizar el error cuadrático medio, el cual se define como:

$$MSE(\hat{\beta}) = E(\hat{\beta} - \beta)^2$$

No es difícil mostrar que esta definición es equivalente a:

$$MSE = \left[Tendencia(\hat{\beta}) \right]^2 + Var(\hat{\beta})$$

1.8.4.4 CONSISTENCIA

La diferencia entre el valor estimado del parámetro y el real se anula para una muestra infinita (es decir para la población).

$$\lim_{n \rightarrow \infty} \hat{\beta}_n = \beta$$

1.9 REGRESIÓN MÚLTIPLE

El análisis de regresión múltiple es una técnica estadística que puede utilizarse para analizar la relación entre una única variable dependiente (o respuesta) y varias variables independientes (o predictores), el objetivo es usar las variables independientes cuyos valores son conocidos para predecir la única variable respuesta seleccionada por el investigador. Cada variable predictor es ponderado

de forma que las ponderaciones indican su contribución relativa a la predicción conjunta. Al calcular las ponderaciones el procedimiento de análisis de regresión asegura la máxima predicción a partir del conjunto de variables independientes. Estas ponderaciones facilitan también interpretación de la influencia en la realización de la predicción.

El modelo:

$$Y = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \dots + \beta_k X_{ki} + \varepsilon_i$$

Donde:

Y es la variable dependiente, las X son las variables independientes y ε_i es el término del error, X_{2i} representa por ejemplo, la i-ésima observación en la variable independiente X_2 . β_1 , es el término constante o intercepto de la ecuación. Las variables independientes son no aleatorias.

Los supuestos de los errores son los siguientes:

$$1. E(e_i) = 0, \forall_i, i = 1, \dots, n$$

$$2. Var(e_i) = \sigma^2, \forall_i, i = 1, \dots, n$$

$$3. Cov(e_i e_j) = 0 \quad \forall_{i,j}, i \neq j, i = 1, \dots, n, j = 1, \dots, n$$

1.9.1 SIGNIFICANCIA DE LA REGRESIÓN

Considere la hipótesis $H_0 : \beta_2 = \beta_3 = \dots = \beta_k = 0$. Esta hipótesis significa que Y no se relaciona con ninguna de las X. La prueba H_0 es llamada prueba de significancia de la regresión. Si las suposiciones del modelo de regresión satisfacen 1, 2, 3 de supuesto y H_0 es verdadero, entonces la razón

$$F = \frac{MSR}{MSE}$$

tiene una distribución F con $df = k, k-1$ grados de libertad. De este modo, la razón F se puede usar para probar la significancia de la regresión.

TABLA 2. ANOVA PARA REGRESIÓN MÚLTIPLE

Fuente	Suma de cuadrados	Grados de libertad df	Cuadrado medio	Razón F
Regresión	SSR	k-1	MSR= SSR/(k-1)	$F = \frac{MSR}{MSE}$
Error	SSE	n - k	MSE= SSE/(n - k)	
Total		n-1		

1.9.2 PRUEBA F PARA LA SIGNIFICANCIA DE LA REGRESIÓN

En el modelo de regresión múltiple, las hipótesis:

$$H_0 : \beta_2 = \beta_3 = \dots = \beta_k = 0$$

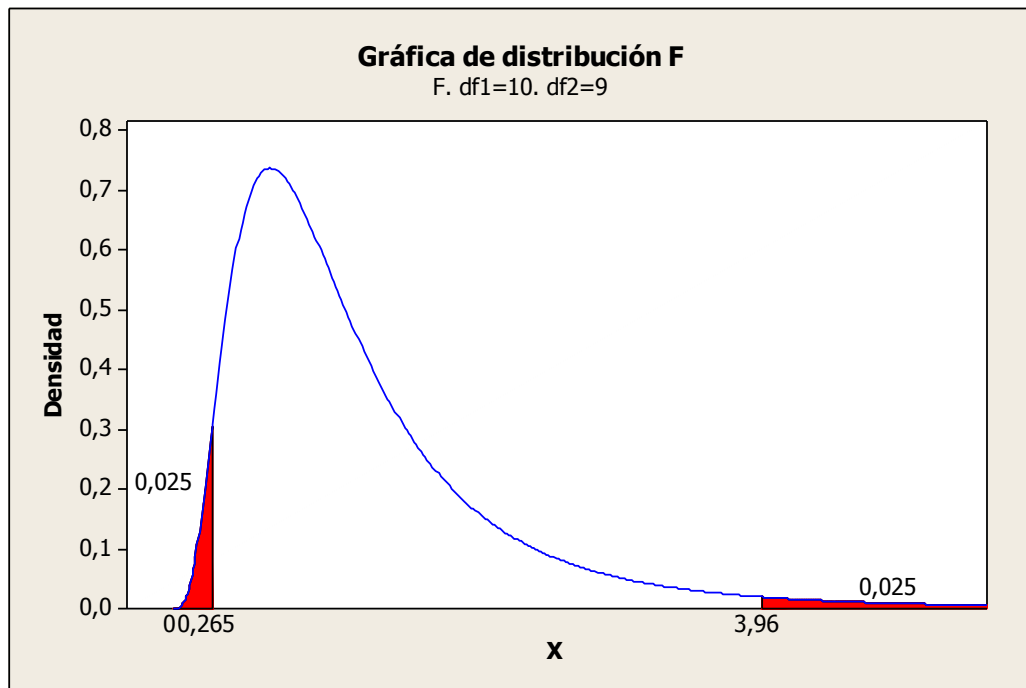
$$H_1 : \exists \beta_j \neq 0, \quad j = 2, \dots, k$$

son probados por la razón

$$F = \frac{MSR}{MSE},$$

con $df = k - 1$ $n - k$ grados de libertad en un nivel de significancia de α , la región de rechazo es

FIGURA 8. GRÁFICA DE DISTRIBUCIÓN F



$$F > F_{\alpha}$$

donde F_{α} es el punto que deja un área α a la derecha en una distribución F con $k - 1, n - k$, grados de libertad del numerador y del denominador respectivamente.

1.9.3 PRUEBAS INDIVIDUALES DEL MODELO BÁSICO DE REGRESIÓN MÚLTIPLE.

En la función de regresión, el coeficiente de una X individual mide el efecto parcial, o neto de esa X en la respuesta Y , manteniendo a las otras X en la ecuación constante. Si la regresión es considerada significativa, es de interés examinar la significancia de las variables explicativas individuales. La pregunta es: dadas las otras X , ¿el efecto de una X particular es importante o se puede eliminar esta X de la función de regresión? Esta pregunta tiene la respuesta al examinar un valor t apropiado.

Para determinar la significancia del término j -ésimo $j = 2, \dots, k$, en la función de regresión, la estadística de prueba t se compara con un punto de la distribución t con grados de libertad $n - k$. Una prueba en el nivel de significancia α de

$$\begin{aligned} H_0 : \beta_j &= 0 \\ H_1 : \beta_j &\neq 0 \end{aligned} \quad j = 2, \dots, k$$

Se rechaza H_0 si $|t| > t_{\alpha/2}$. Aquí $t_{\alpha/2}$ es el punto que deja un área $\alpha/2$ a la derecha en una distribución t con $df = n - k$ grados de libertad. Se debe tener cuidado cuando se eliminen variables explicativas de la función de regresión que son considerados insignificantes por la prueba t ($H_0 : \beta_j = 0$ no puede ser rechazada).

Si las X están relacionadas (caso de multicolinealidad), los coeficientes de los mínimos cuadrados y los valores correspondientes t pueden cambiar, algunas veces considerablemente, si se elimina una sola X de la función de regresión, una X que antes era insignificante puede volverse significativa. En consecuencia, si existe varios valores de t pequeños (insignificantes), las variables explicativas deben ser eliminados una por una (empezando con la variable con el valor t más pequeño) en lugar de un grupo. El proceso se detiene cuando la regresión es significativa y todas las variables tienen estadísticas t grandes (significativas).

TABLA 3. INTERPRETACIÓN DE LOS CONTRASTES

Contraste conjunto (F)	Contraste individual (t)	Diagnóstico
Significativo	Todos significativos	Modelo correcto
Significativo	Algunos significativos	Eliminar variables
Significativo	Todos significativos	Multicolinealidad
No significativo	Todos significativos	Multicolinealidad, variables con sentidos opuestos
No significativo	Algunos significativos	Multicolinealidad, variables con sentidos opuestos
No significativo	Ninguno significativo	Modelo inservible

1.9.4 COEFICIENTE DE CORRELACIÓN

En el análisis de correlación, el objetivo principal es medir la fuerza o el grado de asociación lineal entre dos variables. El coeficiente de correlación, mide esta fuerza de asociación lineal.

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{n \sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

Los valores que puede tomar el coeficiente de correlación r son:

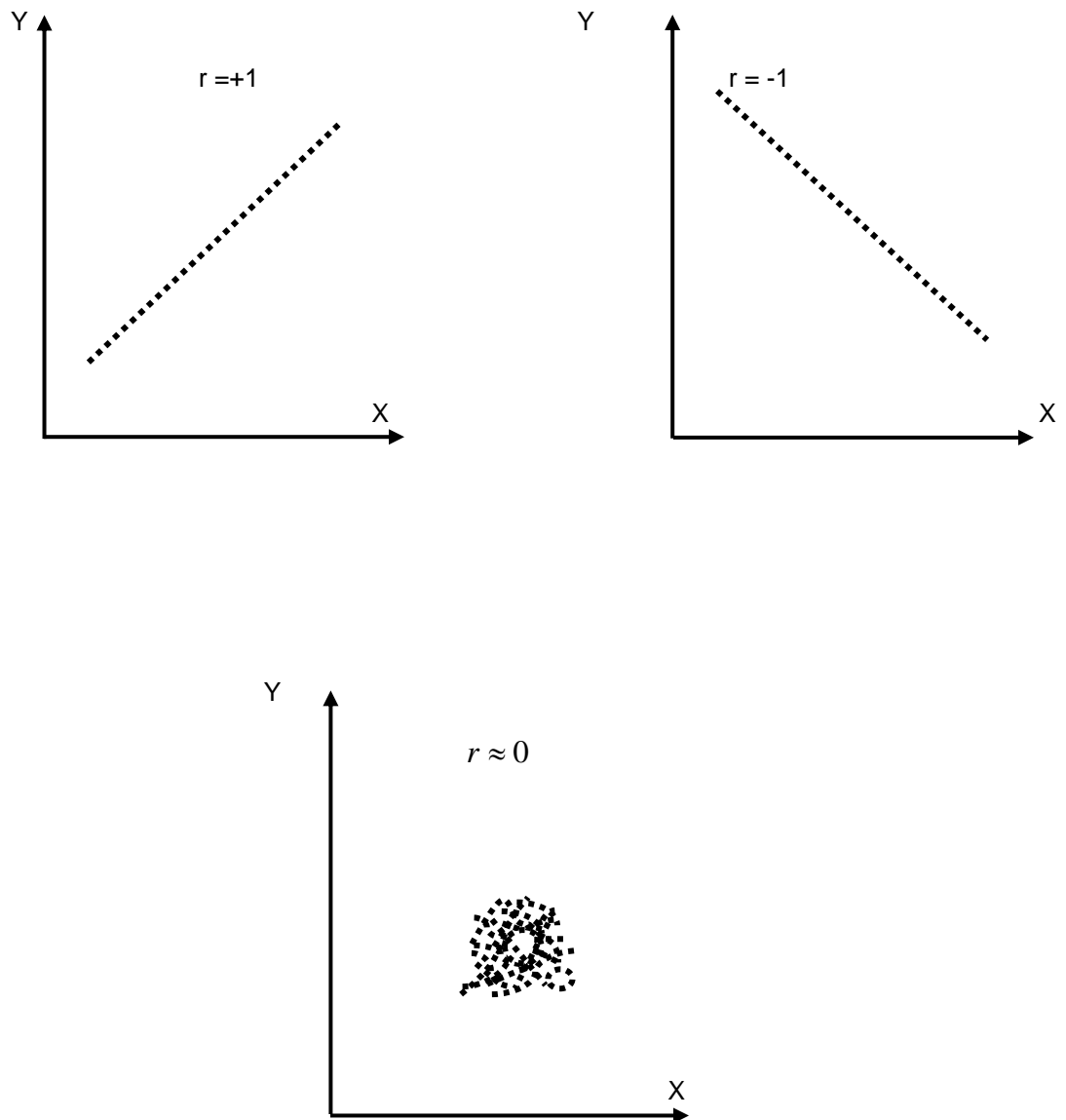
$$-1 \leq r \leq 1$$

Si $r = 0$ no existe relación alguna.

Si $r = 1$ existe relación perfecta positiva.

Si $r = -1$ existe relación perfecta negativa.

FIGURA 9. GRÁFICA DE COEFICIENTE DE CORRELACIÓN



1.9.5 COEFICIENTE DE DETERMINACIÓN

La R^2 mide la proporción de la variación en Y que es explicada por la ecuación de regresión múltiple. R^2 a menudo se usa de manera informal como una estadística de bondad de ajuste y para comparar la validez de los resultados de la regresión bajo especificaciones alternativas de las variables independientes en el modelo. Entonces definimos R^2 como:

$$R^2 = \frac{\sum \left(\hat{Y}_i - \bar{Y} \right)^2}{\sum \left(Y_i - \bar{Y} \right)^2} = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}$$

El coeficiente de determinación se utiliza a veces para comparar la eficacia de distintas regresiones, pero esta utilización no está justificada y representa problemas importantes. En primer lugar, R^2 aumenta siempre introduciendo nuevas variables en el modelo, aunque su efecto no sea significativo. Por lo tanto, siempre se puede artificialmente incrementar su valor añadiendo nuevas variables. En segundo lugar R^2 es muy sensible a la formulación del modelo y la elección de la variable dependiente: es posible estimar dos modelos formalmente idénticos y con la misma capacidad predictiva pero que conducen a distintos valores de R^2

1.9.6 COEFICIENTE DE DETERMINACIÓN AJUSTADO \bar{R}^2

Al igual que R^2 , el coeficiente de determinación ajustado \bar{R}^2 es una medida de la proporción de variabilidad en la respuesta Y explicada por la regresión, el término ajustado significa ajustado por los grados de libertad (df) asociados a la suma de cuadrados que se consideran en la ecuación anterior. El coeficiente ajustado de determinación \bar{R}^2 es dado por:

$$\bar{R}^2 = 1 - (1 - R^2) \frac{n-1}{n-k}$$

Por la ecuación se comprende que:

1. Si $k = 1$, entonces $R^2 = \bar{R}^2$
2. Si $k > 1$, entonces $R^2 \geq \bar{R}^2$.
3. \bar{R}^2 puede ser negativa.

1.10 EXTENSIÓN DEL MODELO DE REGRESIÓN LINEAL MÚLTIPLE:

1.10.1 MULTICOLINEALIDAD

La multicolinealidad hace referencia a la relación lineal entre las variables independientes. Sus efectos sobre la estimación son inciertos dependiendo del grado de linealidad de las variables independientes y de la importancia que estas tengan sobre la variable dependiente. La presencia de multicolinealidad no constituye un problema. El problema es el grado con que las variables independientes están linealmente correlacionadas.

1.10.2 DETECCIÓN DE LA MULTICOLINEALIDAD

Una R^2 relativamente alta en una ecuación con una estadística t poco significativa es un indicador de la multicolinealidad. De hecho, es posible que la estadística F para la ecuación de regresión sea altamente significativa, mientras que ninguna de las estadísticas t individuales sea significativa en sí misma.

1.10.2.1 ALGUNAS POSIBLES CORRECCIONES A LA MULTICOLINEALIDAD

Cuando se presenta la multicolinealidad, se acude a diferentes soluciones propuestas aunque ninguna de ellas suele resultar plenamente satisfactoria y, en algunas ocasiones, no es posible llevarlas a la práctica. Dejando aparte las

soluciones que requieren métodos estadísticos más avanzados, las medidas a tomar pueden concretarse en los puntos siguientes:

1.10.2.2 ELIMINAR VARIABLES

El problema de multicolinealidad se reduce, de forma esencial, a una falta de información suficiente en la muestra que permita estimar de forma más o menos exacta los parámetros individuales. En algunas ocasiones podemos estar interesados en unos parámetros más que en otros y, si existe un elevado grado de multicolinealidad entre las variables independientes que forman parte de la regresión, podemos tratar de excluir una de ellas y estimar los parámetros correspondientes al resto.

1.10.2.3 AUMENTAR EL TAMAÑO MUESTRAL

Esta es una solución que a menudo se suele sugerir pero que, en numerosas ocasiones no es factible (porque para la estimación del modelo seguramente se habrá considerado todas las observaciones disponibles para las variables) y que en otras, no conduce a una eliminación o reducción del problema (porque lo relevante no es el número de observaciones, sino el contenido informativo de las

mismas) al añadir nuevas observaciones, en las mismas se sigue cumpliendo el mismo patrón de multicolinealidad.

1.10. 2.4 TRANSFORMACIÓN DE LOS DATOS

Las transformaciones pueden aplicarse tanto a las variables dependientes como a independientes como o a ambas. El tipo específico y necesario de transformaciones puede basarse en razones teóricas (tales como transformar una relación no lineal conocida) o por razones empíricas (identificadas por medios gráficos o estadísticos).

1.10.3 HETEROCEDASTICIDAD

El problema de heterocedasticidad se presenta cuando la varianza de los errores no es constante para diferentes observaciones. La heterocedasticidad tiene que ver con la relación entre una o más de las variables independientes del modelo y el cuadrado de los errores estimados a partir de la regresión. Este problema se manifiesta en un crecimiento o decrecimiento de la varianza del modelo.

$$Var(e_i) = \sigma_i^2, \quad i = 2, \dots, k$$

1.10.4 DETECCIÓN DE LA HETEROCEDASTICIDAD

1.10.4.1 MÉTODO GRÁFICO

Un test para detectar la presencia de heterocedasticidad consiste simplemente en analizar cuidadosamente la gráfica de los residuos e_i . De la inspección de dicho gráfico se puede obtener una idea clara sobre la presencia de una conducta sistemática de los residuos.

1.10.4.2 LA PRUEBA DE WHITE

Hal White propuso una prueba estrechamente relacionada que no depende de la normalidad. Usamos los residuales de la regresión para ejecutar la siguiente regresión:

$$\hat{\varepsilon}_i^2 = \alpha_1 + \alpha_2 X_i + v_i$$

a partir de la cual calculamos la medida de bondad de ajuste R^2 . La prueba de White se basa en el hecho de que cuando hay heterocedasticidad.

$$nR^2 \sim \chi^2$$

Las hipótesis nula y alternativa son:

H_0 : No hay heterocedasticidad

H_1 : Hay heterocedasticidad

Según la hipótesis nula de que no hay heterocedasticidad, puede demostrarse que el tamaño de la muestra (n) multiplicado por R^2 obtenido de la regresión auxiliar asintóticamente sigue la distribución ji cuadrada, con grados de libertad igual al número de variables independientes (sin el término constante) en la regresión auxiliar.

1.11. ANALISIS DE RESIDUALES

La información sobre la variación que no se puede explicar por la función de regresión ajustada encuentra en los residuos $e_i = Y_i - \hat{Y}_i$. Para verificar los méritos de un modelo tentativo, se pueden examinar diversas gráficas de residuales.

1. Histograma de los residuales.
2. Gráfica de dispersión, residuales contra los valores ajustados.
3. Gráfica de dispersión, residuales contra la variable explicativa.
4. Gráfica de residuales contra el tiempo si los datos son cronológicos.

1.12 ELIMINACIÓN PROGRESIVA

Este método comienza con una regresión que incorpora todas las variables independientes potencialmente influyentes. A continuación, se calcula los estadísticos t para cada coeficiente, y si alguno de estos valores es menor que un valor prefijado, se elimina esta variable. Se calcula la regresión con las $k - 1$ variables restantes, y se repite el procedimiento de eliminación de variables no significativas.

La estrategia de eliminación progresiva tiene el inconveniente de utilizar mucha capacidad de cálculo, es posible que únicamente un subconjunto pequeño de las k variables sea significativo, y este procedimiento obliga a efectuar regresiones muy extensas. Además, conduce fácilmente al problema de multicolinealidad si hay variables muy relacionadas o k es muy elevado. En contrapartida, es excelente para evitar la exclusión el número de variables es pequeño.

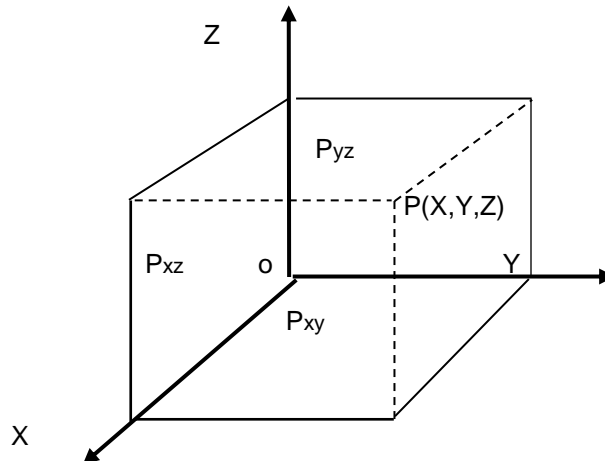
1.13 ESPACIO TRIDIMENSIONAL

Llamaremos espacio tridimensional al conjunto de todas las ternas ordenadas de números reales, donde cada terna ordenada (x, y, z) se denomina punto del espacio tridimensional.

1.13.1 EJES COORDENADOS

Los ejes coordenados son generalmente identificados por las letras X, Y, Z y hablamos frecuentemente del eje X, del eje Y y del eje Z.

FIGURA 10. ESPACIO TRIDIMENSIONAL



La dirección positiva se indica por medio de una flecha, los ejes de coordenadas tomados de dos en dos determinan tres planos llamados planos coordenados: plano XY, plano XZ, y plano YZ.

CAPITULO II

2. MARCO METODOLÓGICO

2.1 PROCEDIMIENTOS PARA ANÁLISIS DE DATOS Y ELABORACIÓN DEL MODELO

A continuación se presenta los procedimientos utilizados en la elaboración del presente trabajo de investigación:

2.2. Elección de variables coherentes y tabulación de los datos de los archivos del departamento forestal del ministerio del ambiente de Chimborazo.

2.3 APLICACIÓN DE A.E.D.

Para realizar un A.E.D. conviene seguir las siguientes etapas:

2.3.1 Preparar los datos para hacerlos accesibles a cualquier técnica estadística.

2.3.2 Realizar un examen gráfico de la naturaleza de las variables individuales a analizar y un análisis descriptivo numérico que permita cuantificar algunos aspectos gráficos de los datos.

2.3.3 Realizar un examen gráfico de las relaciones entre las variables analizadas y un análisis descriptivo numérico que cuantifique el grado de interrelación existente entre ellas.

2.3.4 Evaluar, si fuera necesario, algunos supuestos básicos subyacentes a muchas técnicas estadísticas como, por ejemplo, la normalidad, linealidad y homocedasticidad.

2.3.5 Identificar los posibles casos atípicos y evaluar el impacto potencial que puedan ejercer en análisis estadísticos posteriores.

2.3.6 Evaluar, si fuera necesario, el impacto potencial que pueden tener los datos ausentes sobre la representatividad de los datos analizados.

2.4 SELECCIÓN DE LAS VARIABLES INCLUIDAS EN UNA REGRESIÓN MÚLTIPLE.

Para realizar una regresión múltiple es conveniente seguir los siguientes pasos:

2.4.1 Calcular la matriz de correlación para obtener una visión global de las interrelaciones existentes entre las mismas.

2.4.2. El valor de R^2 debe ser lo suficiente grande, si se trata de describir el comportamiento de una variable independiente. Entre modelos alternativos se elegirá aquel que, cumpliendo las reglas anteriores, alcancen un mayor \overline{R}^2 .

2.4.3. El estadístico F y los estadísticos t deben rechazar las hipótesis nulas que cada test contrasta. No obstante, pueden darse ciertos casos en los que siendo significativas conjuntamente, no lo sea individualmente. Aun así el modelo estimado puede utilizarse, por ejemplo, para realizarse predicciones. Ver tabla de la interpretación de los contrastes.

2.4.4 El comportamiento de los residuos debe ser aleatorio. El análisis de los residuos, desde el punto de vista gráfico, permitirá intuir si el comportamiento de los mismos es aleatorio, hipótesis que condiciona la construcción de los estadísticos.

2.4.5. Se efectuó la comprobación de la homocedasticidad, en forma gráfica y con la prueba de White.

2.5 PRONÓSTICOS

2.5.1 Evaluación con las medidas de precisión.

2.6. HIPÓTESIS ESTADÍSTICO

2.6.1 Planteamiento de la prueba de hipótesis

H0: El volumen de aprovechamiento de madera de pino será mayor o igual en los próximos meses respecto al promedio mensual.

H1: El volumen de aprovechamiento de madera de pino disminuirá en los próximos meses respecto al promedio mensual.

2.6.2 Especificación del nivel de significancia ($\alpha = 0.05$)

2.6.3 Tamaño de la muestra y cálculo de estadísticos.

2.6.4 Establecimiento de RA Y RR.

2.6.5 Decisión estadística

CAPÍTULO III

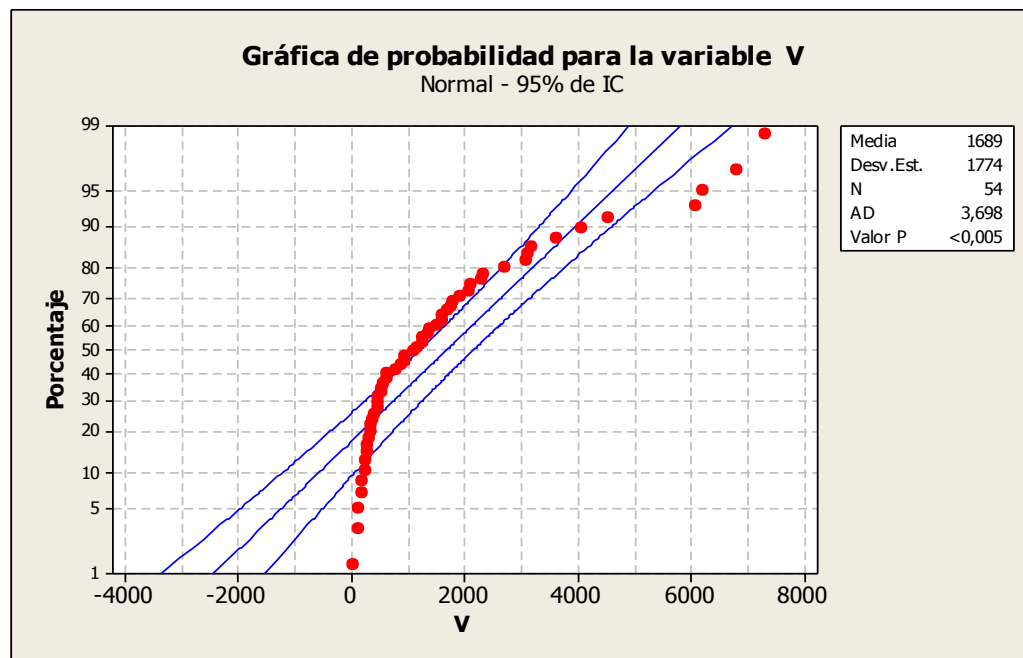
3. ANÁLISIS DE RESULTADOS

Se incorporó los datos disponibles para las 6 variables, (promedios mensuales de 4 años y medio, a partir del 2008 a 2012)

Para lo anterior tenemos los datos a partir de enero del 2008 a junio del 2012 es decir 54 observaciones.

3.1 TEST DE NORMALIDAD

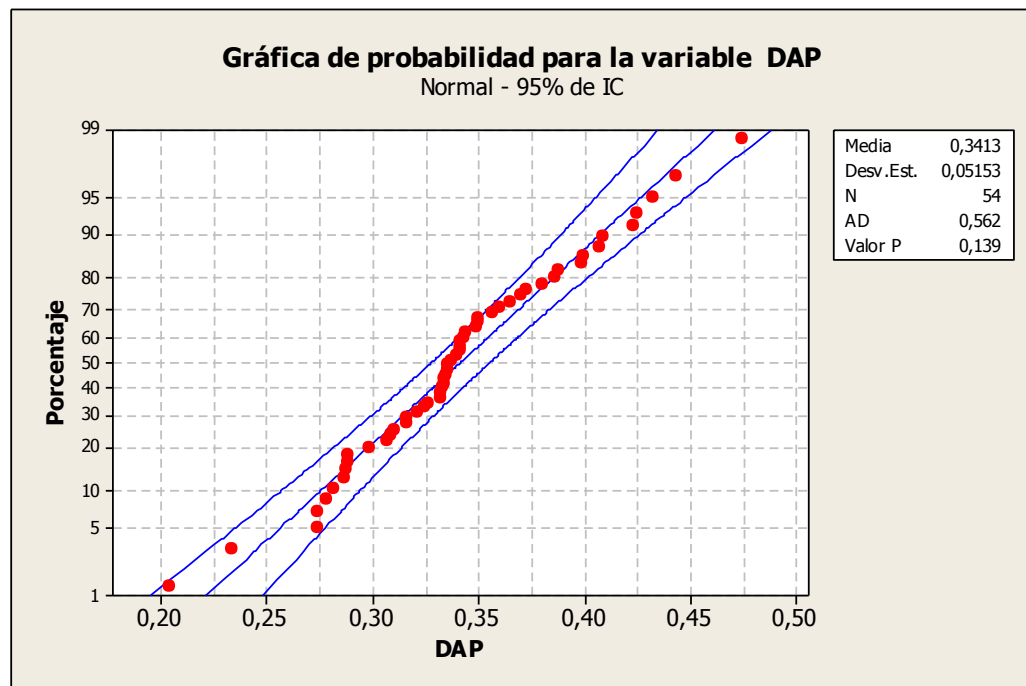
FIGURA 11. GRÁFICA NORMAL DE PROBABILIDAD



Elaborado por: Tesista
Fuente: MAE-Chimborazo

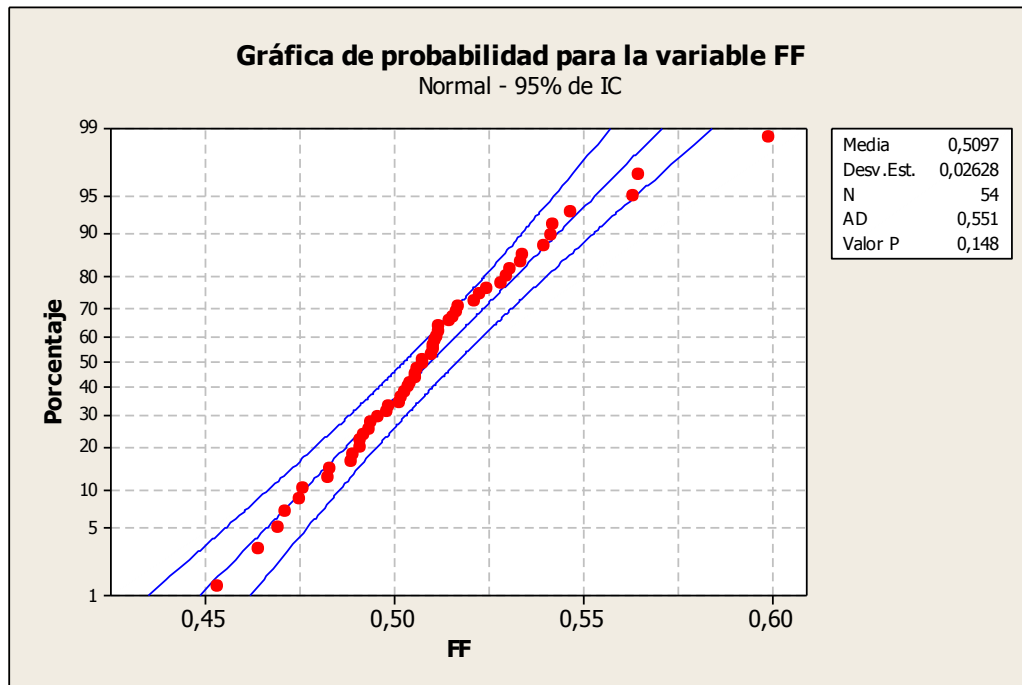
Según la figura 11, para variable volumen los puntos están fuera de los límites de confianza al 95% además el valor de p es menor a 0,05 se acepta la hipótesis alternativa, por tanto la variable volumen no proviene de una distribución normal.

FIGURA 12. GRÁFICA NORMAL DE PROBABILIDAD



Elaborado por: Tesista
Fuente: MAE-Chimborazo

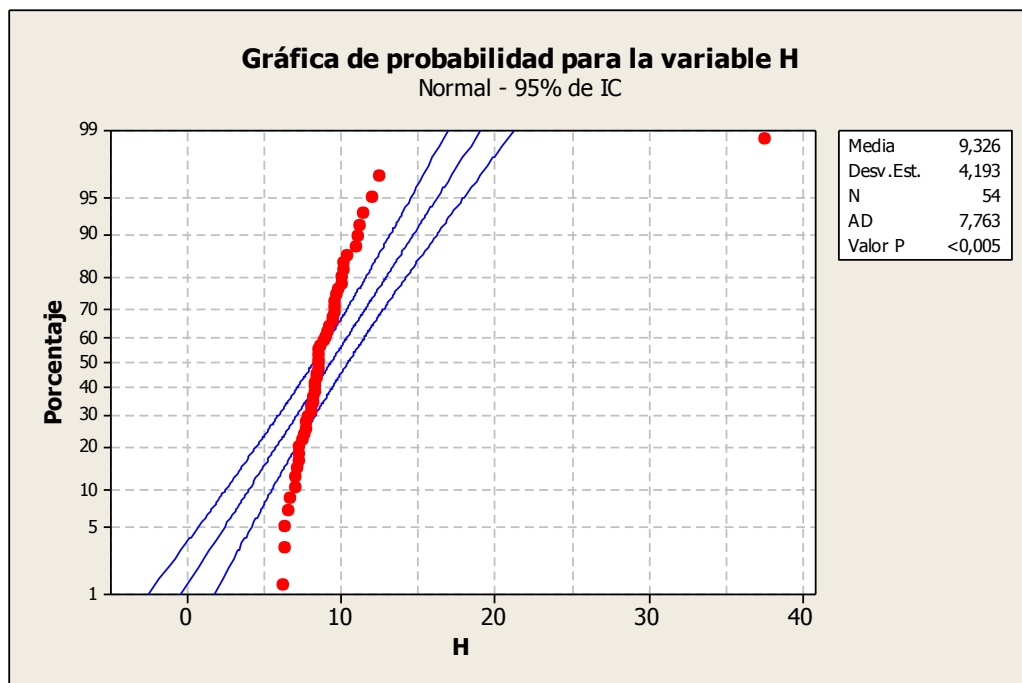
FIGURA 13. GRÁFICA NORMAL DE PROBABILIDAD



Elaborado por: Tesista

Fuente: MAE-Chimborazo

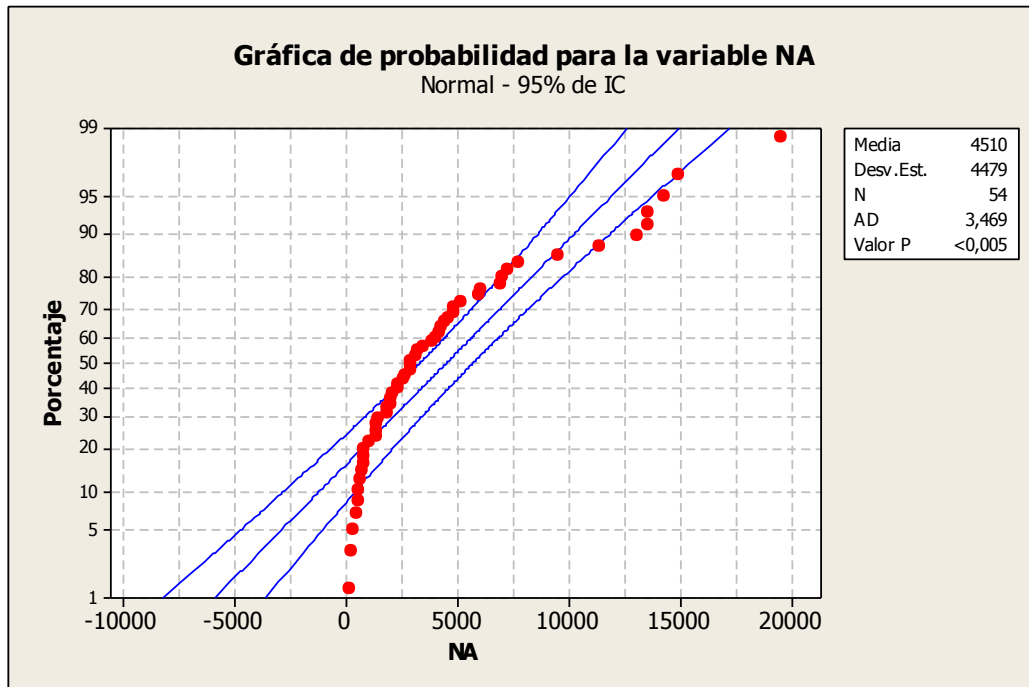
FIGURA 14. GRÁFICA NORMAL DE PROBABILIDAD



Elaborado por: Tesista

Fuente: MAE-Chimborazo

FIGURA 15. GRÁFICA NORMAL DE PROBABILIDAD



Elaborado por: Tesista

Fuente: MAE-Chimborazo

Obsérvese en la figura 12, que todos los puntos están dentro de los límites de confianza al 95% y el valor p es 0,139 que es mayor 0,05 por lo que se acepta la hipótesis nula, por ende la variable diámetro a la altura del pecho DAP, proviene de una distribución normal

En la figura 13, se puede observar que los puntos están dentro de los límites de confianza al 95% y el valor p es 0,148 que es mayor 0,05 por lo que se acepta la hipótesis nula, consecuentemente la variable factor de forma FF, proviene de una distribución normal.

En la figura 14, se puede apreciar para la variable altura H, la mayoría de los puntos están fuera de los límites de confianza al 95% y el valor p es menor al 0,05 entonces se acepta la hipótesis alternativa, por ende que la variable altura H, no proviene de una distribución normal

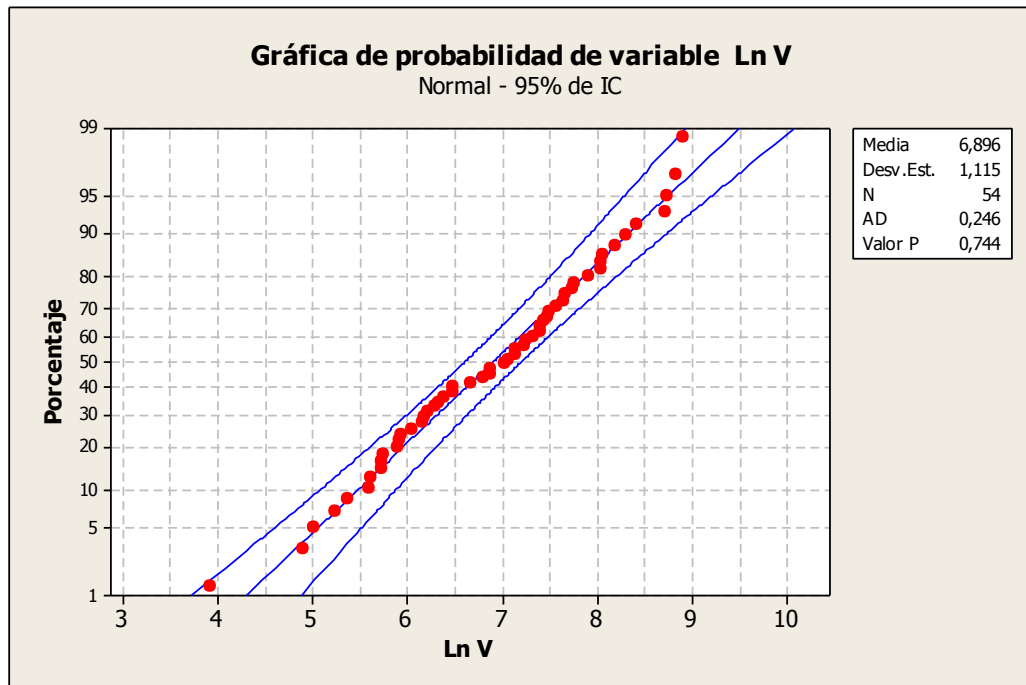
Obsérvese en la figura 15, la mayoría de los puntos están fuera de los límites de confianza al 95% y el valor p es menor al 0,05 entonces se acepta la hipótesis alternativa, entonces concluimos que variable números de árboles NA, no proviene de una distribución normal

3.2 TRANSFORMACIÓN DE DATOS

Las transformaciones nos permiten disminuir la dispersión y conseguir linealidad entre las variables, transformamos con logaritmo natural (\ln) a todas las variables en estudio.

Nota: nótese que tan solo a las variables que no cumplían la normalidad realizaremos nuevamente el test de normalidad, mientras que utilizaremos todas las variables transformadas para la regresión.

FIGURA 16. GRÁFICA NORMAL DE PROBABILIDAD

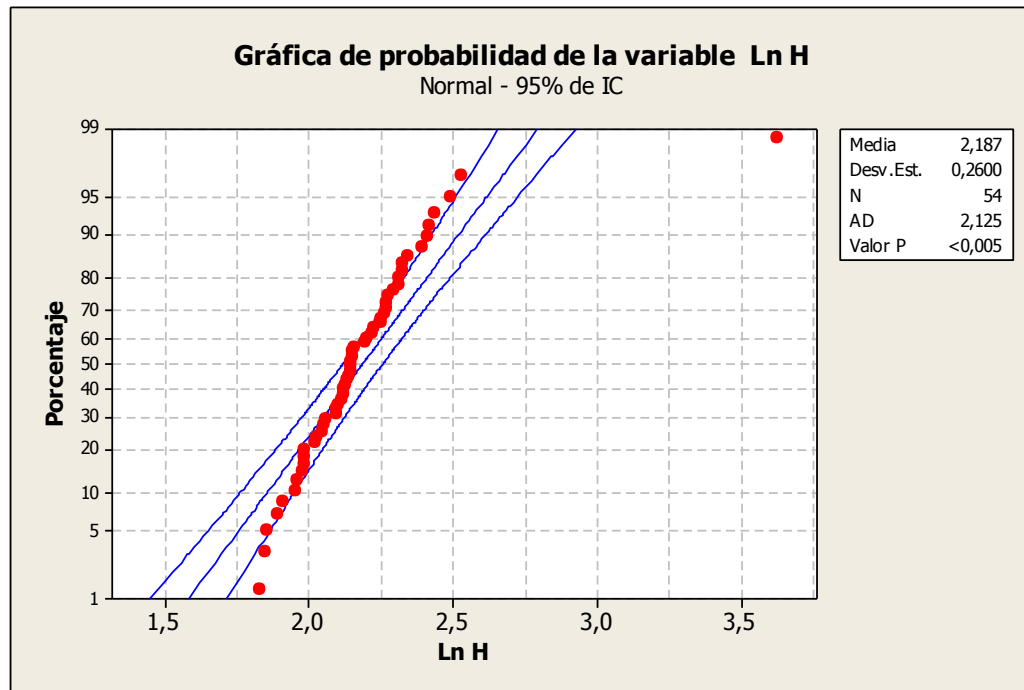


Elaborado por: Tesista

Fuente: MAE-Chimborazo

En la figura 16, con los datos transformados, se puede observar que los puntos están dentro de los límites de confianza al 95% y el valor p es 0,744 que es mayor 0,05 por lo que se acepta la hipótesis nula, consecuentemente la variable volumen V , proviene de una distribución normal.

FIGURA 17. GRÁFICA NORMAL DE PROBABILIDAD

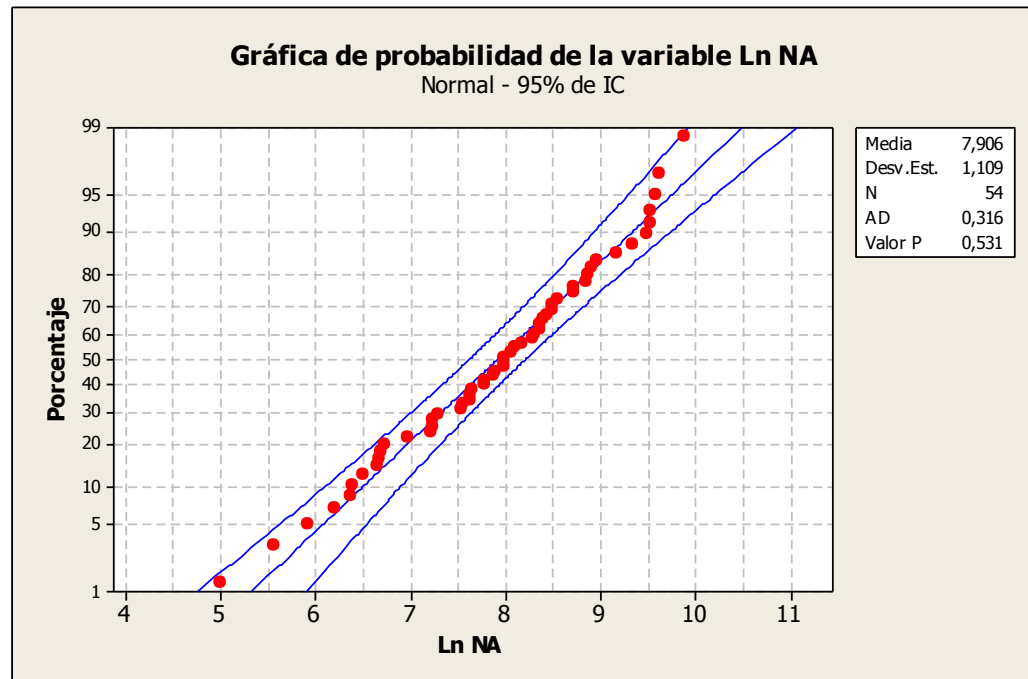


Elaborado por: Tesista

Fuente: MAE-Chimborazo

En la figura 17, se puede apreciar que los puntos están fuera de los límites de confianza al 95% y el valor p es menor al 0,05 entonces se acepta la hipótesis alternativa, por tanto la variable altura H, no proviene de una distribución normal.

FIGURA 18. GRÁFICA NORMAL DE PROBABILIDAD



Elaborado por: Tesista

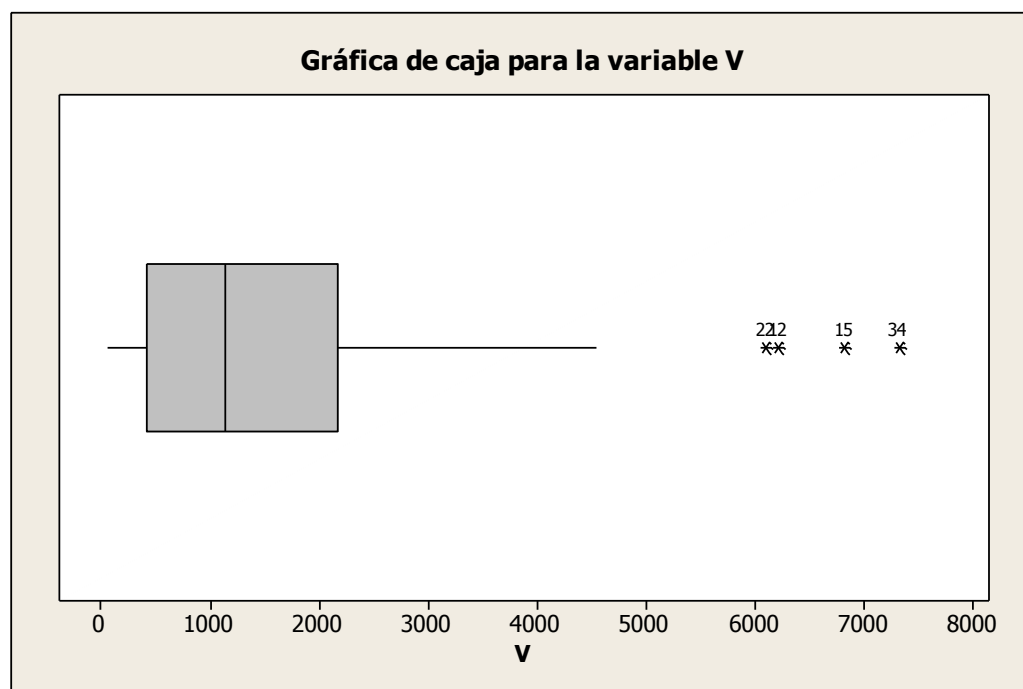
Fuente: MAE-Chimborazo

Obsérvese en la figura 18, que todos los puntos están dentro de los límites de confianza al 95% y el valor p es 0,531 que es mayor 0,05 por lo que se acepta la hipótesis nula, consecuentemente la variable número de árboles NA, provienen de una distribución normal.

Vamos a seleccionar las variables que más ayudan al modelo para pronosticar, también observar la existencia de datos atípicos, ya que nos pueden alterar el modelo.

Se realizará representación gráfica en el diagrama de caja para observar datos atípicos y ver la forma de las distribuciones de las variables en estudio.

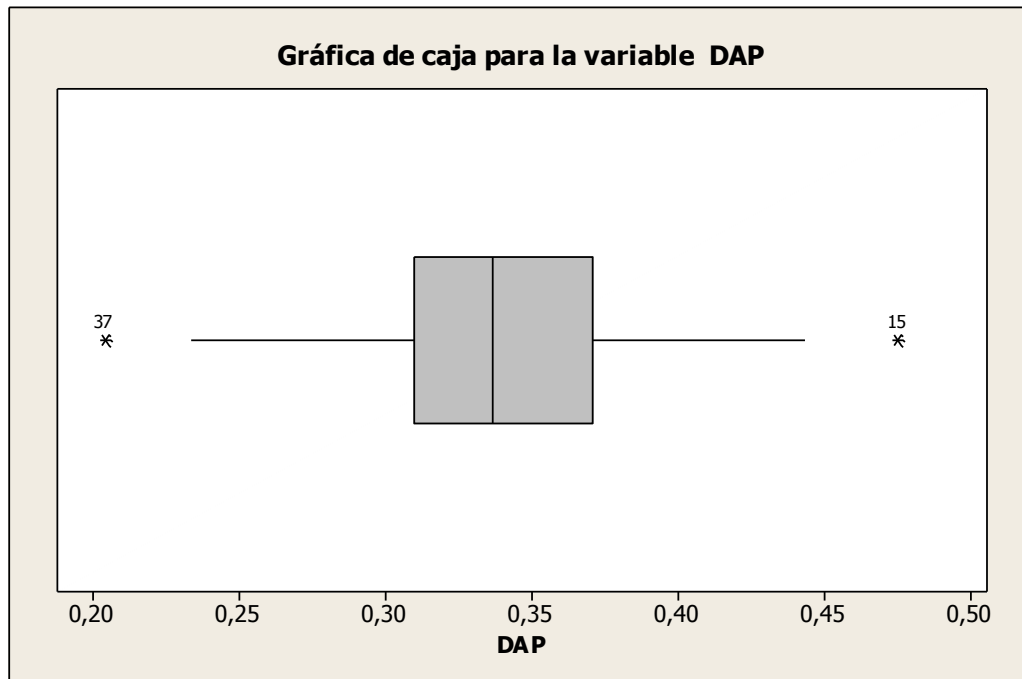
FIGURA 19. DIAGRAMA DE CAJA



Elaborado por: Tesista
Fuente: MAE-Chimborazo

Según la figura 19, se puede apreciar que existe 4 datos atípicos los cuales son los datos 22, 12, 15, 34, de la variable volumen de madera medida en m^3 . La distribución es asimétrica hacia la derecha, el sesgo es positivo.

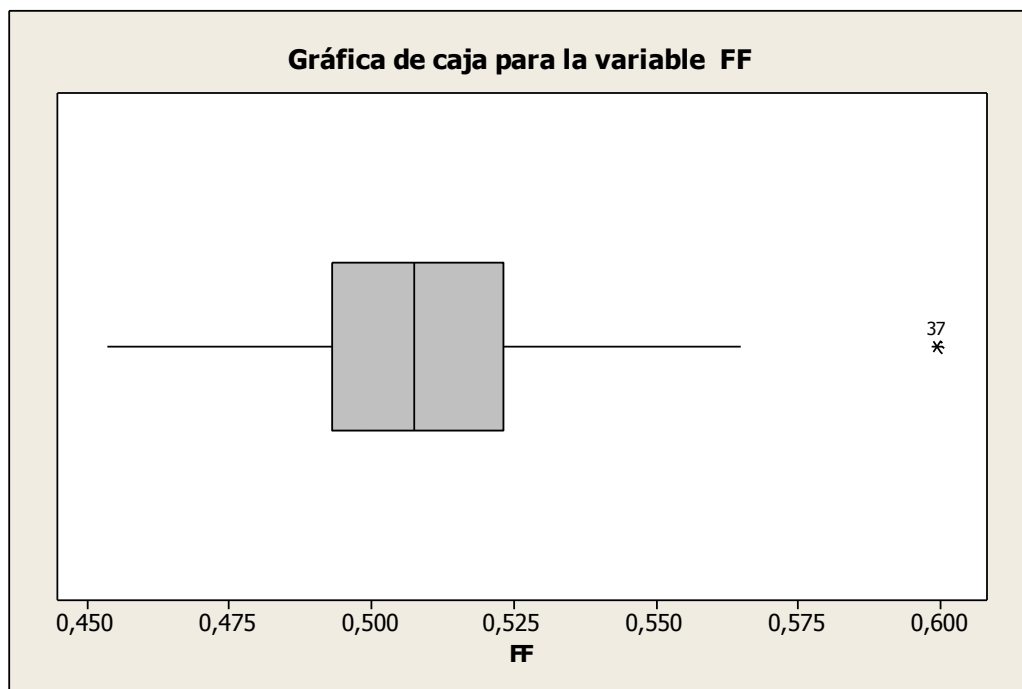
FIGURA 20. DIAGRAMA DE CAJA



Elaborado por: Tesista

Fuente: MAE-Chimborazo

FIGURA 21. DIAGRAMA DE CAJA



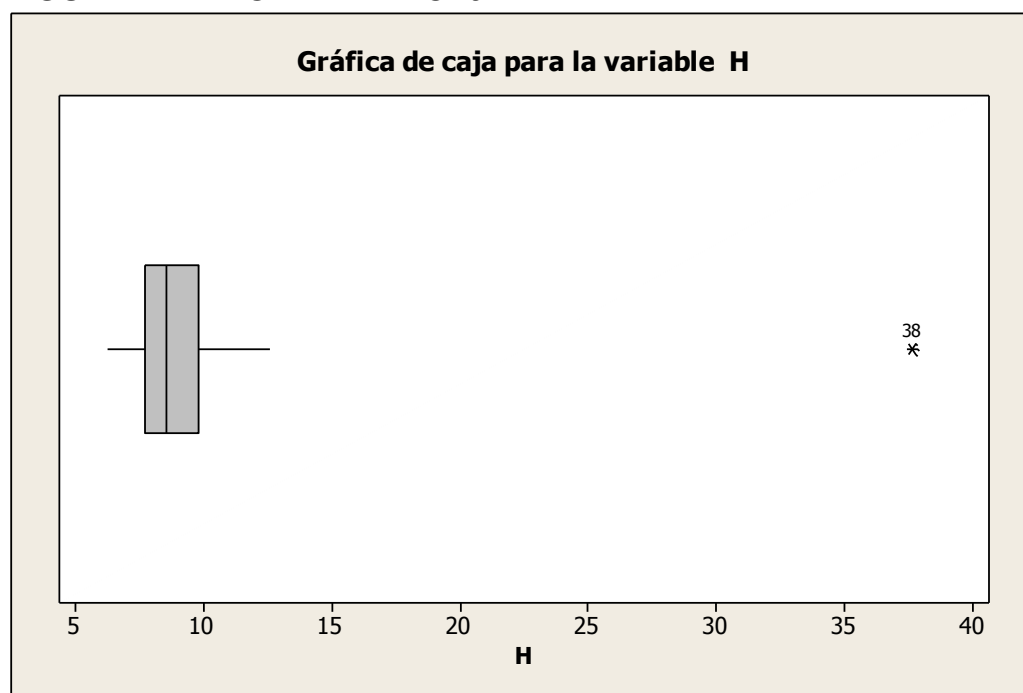
Elaborado por: Tesista

Fuente: MAE-Chimborazo

Se observa en la figura 20, que existe 2 datos atípicos para la variable DAP, son los datos 15 y 37 y es asimétrica, tiene sesgo positivo, hay mayor concentración de valores a la derecha.

Se observe en la figura 21, que existe un dato atípico para la variable FF, que corresponde al dato 37, es asimétrica, hay mayor valor concentrado a la derecha de la media que a la de su izquierda, tiene sesgo positivo.

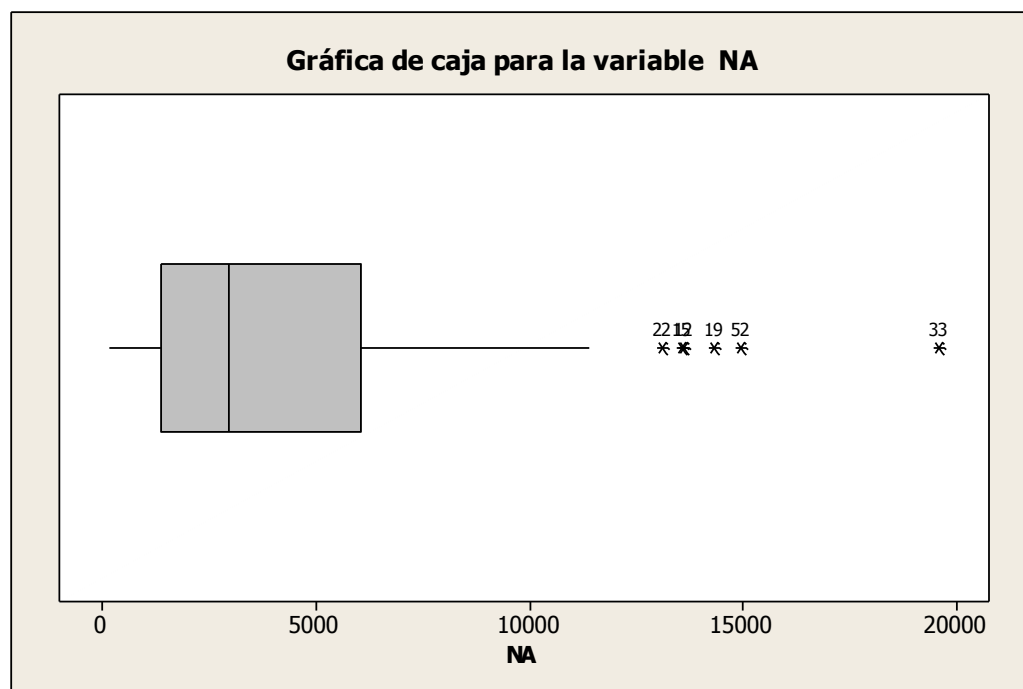
FIGURA 22. DIAGRAMA DE CAJA



Elaborado por: Tesista
Fuente: MAE-Chimborazo

En la figura 22, para la variable altura H tenemos un dato atípico relevante que corresponde al dato 38, la distribución es asimétrica a la derecha y tiene sesgo positivo.

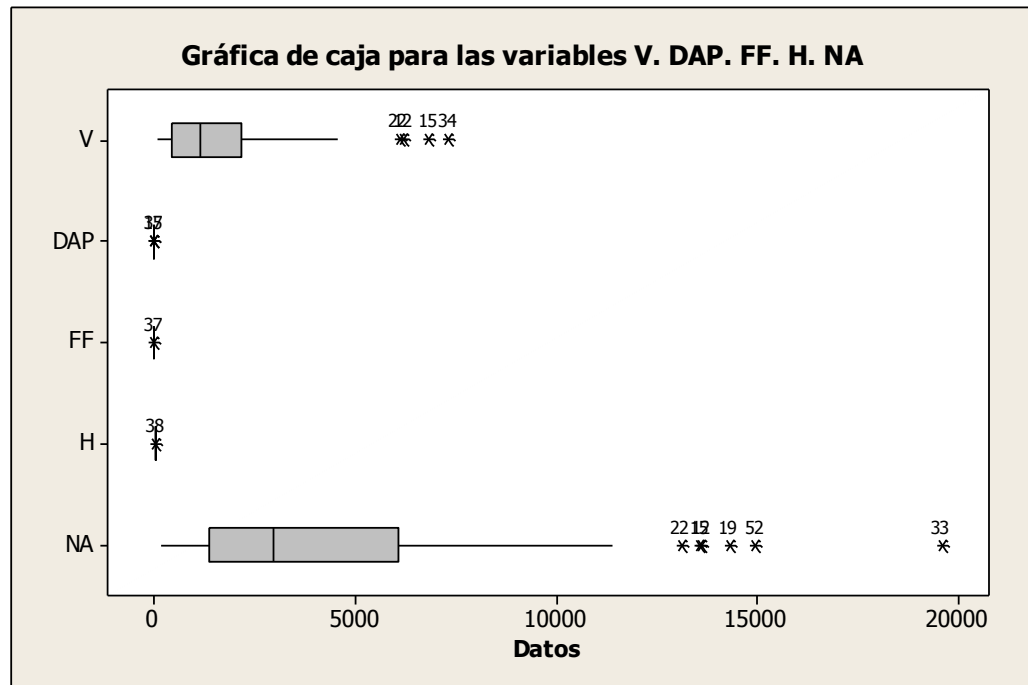
FIGURA 23. DIAGRAMA DE CAJA



Elaborado por: Tesista
Fuente: MAE-Chimborazo

Podemos ver en la figura 23, correspondiente para la variable número de árboles NA, 5 datos atípicos que corresponde a los datos 33, 52, 19, 12, 15, 22, la distribución es asimétrica a la derecha los valores más se concentra al lado derecho de la media, tiene un sesgo positivo.

FIGURA 24. DIAGRAMA DE CAJA



Elaborado por: Tesista

Fuente: MAE-Chimborazo

Observación: son atípicos, los datos 12, en las variables V y NA, el dato 15 en las variables V, DAP, y NA en cambio el dato 22 en las variables V y NA, por último tenemos al dato 37 en las variables DAP y FF.

Para realizar el análisis descriptivo de cada una de las variables que aportan significativamente al modelo, separamos los 9 datos atípicos, en nuestro caso ahora serán 45 datos.

3.3 ANÁLISIS DESCRIPTIVO DE LAS VARIABLES SIGNIFICATIVAS

TABLA 4. ESTADÍSTICAS DESCRIPTIVAS PARA VOLUMEN V

Volumen V	
Media	1129,55
Error típico	140,60
Mediana	896,41
Moda	#N/A
Desviación estándar	943,18
Curtosis	1,58
Coeficiente de asimetría	1,29
Rango	4020,85
Mínimo	50,67
Máximo	4071,52

El volumen promedio mensual de aprovechamiento de madera es $1129,54 \text{ m}^3$, siendo el máximo $4071,51 \text{ m}^3$, mientras que el valor mínimo es $50,67 \text{ m}^3$, tiene una dispersión de 943,18 respecto a la media, según el coeficiente de asimetría de 1,29 indica que la distribución es asimétrica a la derecha existe mayor concentración de los valores a la derecha y tiene sesgo positivo, por el valor del coeficiente de curtosis de 1,58 la distribución es leptocúrtica.

TABLA 5. ESTADÍSTICAS DESCRIPTIVAS PARA DIÁMETRO A LA ALTURA DEL PECHO DAP

Diámetro a la altura del pecho	<i>DAP</i>
Media	0,35
Error típico	0,01
Mediana	0,34
Moda	0,33
Desviación estándar	0,04
Curtosis	-0,08
Coefficiente de asimetría	0,50
Rango	0,17
Mínimo	0,28
Máximo	0,44

El DAP promedio es de 0.35 m, el valor máximo es de 0,44 m, en cambio el mínimo es de 0,28 m, el valor que más se repite o la moda es 0,33m, tiene una desviación estándar, es decir tiene una dispersión de 0,04 respecto al promedio, según el valor del coeficiente de curtosis de -0,08 la distribución es platicúrtica, en cambio por el coeficiente de asimetría de 0,50 indica que la distribución es asimétrica a la derecha tiene sesgo positivo.

TABLA 6. ESTADÍSTICAS DESCRIPTIVAS PARA NÚMERO DE ÁRBOLES NA

Número de árboles	NA
Media	3010,20
Error típico	357,77
Mediana	2629,83
Moda	#N/A
Desviación estándar	2400,02
Curtosis	2,08
Coefficiente de asimetría	1,27
Rango	11250,33
Mínimo	149,00
Máximo	11399,33

La media promedio mensual de árboles es de 3010, tiene como un máximo 11399 árboles, el valor mínimo de árboles es de 149, tiene una dispersión 2400 respecto a la media, tiene un coeficiente de curtosis 2,08 por lo cual la distribución es leptocúrtica y el coeficiente de asimetría 1,27 indica que es asimétrica a la derecha y tiene sesgo positivo.

3.4 ANÁLISIS DESCRIPTIVO DE LOS DATOS ATÍPICOS

TABLA 7. ESTADÍSTICAS DESCRIPTIVAS PARA VOLUMEN V

Volumen V	
Media	4485,08
Error típico	774,36
Mediana	4546,36
Moda	#N/A
Desviación estándar	2323,09
Curtosis	-0,62
Coeficiente de asimetría	-0,49
Rango	7011,41
Mínimo	309,32
Máximo	7320,73

El promedio mensual de volumen es de 4485,08 tiene una dispersión de 2323,09 con respecto a la media, el volumen mensual más alto es 7320,73, el volumen mensual más bajo es 309.32, por el valor de coeficiente de asimetría de -0,49 la distribución es asimétrica a la izquierda tiene sesgo negativo, por el coeficiente de curtosis de -0,62 la distribución es platicúrtica.

TABLA 8. ESTADÍSTICAS DESCRIPTIVAS PARA DIÁMETRO A LA ALTURA DEL PECHO DAP

<i>Diámetro a la altura del pecho</i>	<i>DAP</i>
Media	0,31
Error típico	0,03
Mediana	0,29
Moda	#N/A
Desviación estándar	0,09
Curtosis	0,17
Coefficiente de asimetría	0,91
Rango	0,27
Mínimo	0,20
Máximo	0,48

La media mensual del DAP es de 0,31m, tiene una dispersión de 0,09 m el DAP más alto es 0,48 m, en cambio el DAP más bajo es de 0,20 m, por el coeficiente de asimetría es asimétrica hacia la derecha por ende tiene sesgo positivo, por el valor de coeficiente de curtosis la distribución es leptocúrtica.

TABLA 9. ESTADÍSTICAS DESCRIPTIVAS PARA NÚMERO DE ÁRBOLES NA

Número de árboles	NA
Media	12009,17
Error típico	1676,75
Mediana	13588,00
Moda	#N/A
Desviación estándar	5030,25
Curtosis	0,76
Coefficiente de asimetría	-0,70
Rango	17222,00
Mínimo	2372,20
Máximo	19594,20

La media para el número de árboles mensuales es de 12009, tiene una dispersión de 5030,25 según los valores de coeficiente de curtosis en la tabla anterior la distribución es leptocúrtica, y por el coeficiente de asimetría de -0,70 es asimétrica hacia la izquierda por lo que tiene sesgo negativo.

3.5 ANÁLISIS DE REGRESIÓN

3.5.1 ESTUDIO DE LA MATRIZ DE CORRELACIÓN

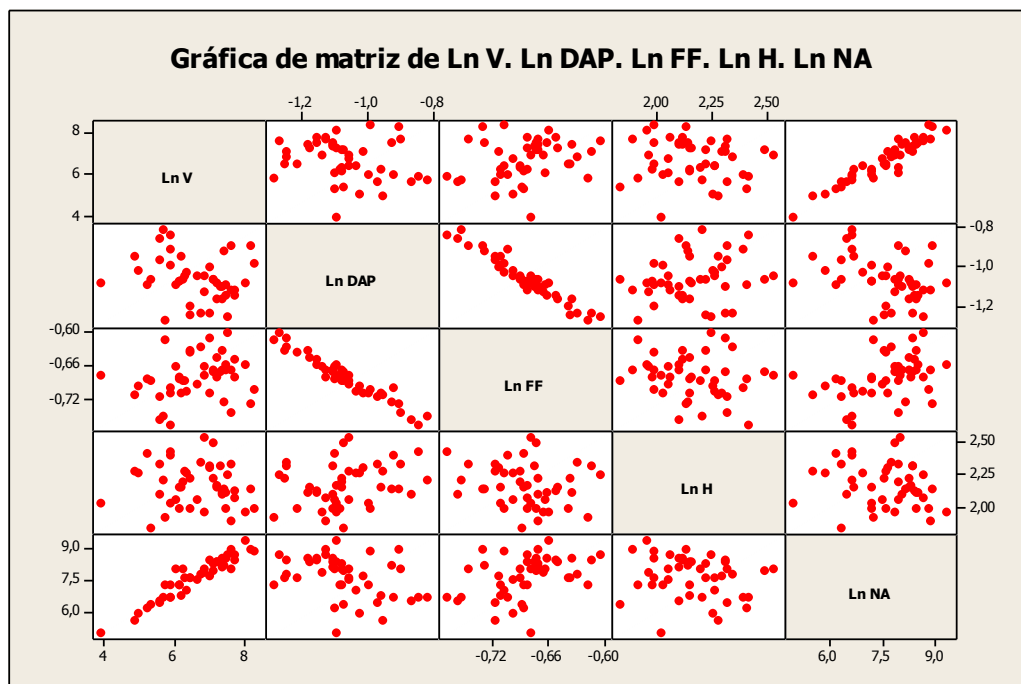
TABLA 10. MATRIZ DE CORRELACIÓN

	<i>Ln V</i>	<i>Ln DAP</i>	<i>Ln FF</i>	<i>Ln H</i>	<i>Ln NA</i>
<i>Ln V</i>	1				
<i>Ln DAP</i>	-0,201	1			
<i>Ln FF</i>	0,234	-0,969	1		
<i>Ln H</i>	-0,088	0,227	-0,184	1	
<i>Ln NA</i>	0,949	-0,356	0,392	-0,173	1

Con el resultado de la matriz de correlación, la variable que más aportaría al modelo es *Ln NA* indica que tiene una correlación positiva de 0,949, la segunda variable sería *Ln FF*, que tiene una correlación moderada de 0,234, el tercero sería *Ln DAP* que indica una correlación moderada negativa de -0,201 y por último *Ln H* que indica una moderada correlación negativa -0.088, la misma que no aporta al modelo.

Se puede notar que existe multicolinealidad entre las variables Ln DAP y Ln FF están sumamente intercorrelacionadas, por lo que la variable Ln FF debe ser eliminado.

FIGURA 25. DIAGRAMA DE DISPERSIÓN



Elaborado por: Tesista
Fuente: MAE-Chimborazo

Podemos observar según figura 19, que la variable Ln DAP y Ln FF gráficamente existe una relación fuerte negativa, están muy correlacionados. Además muestra que Ln H no está correlacionado con ninguno de las otras cuatro variables.

3.5.2 RESULTADOS REGRESIÓN MÚLTIPLE

TABLA 11. RESULTADOS DE LA REGRESIÓN MÚLTIPLE

Estadísticas de la regresión	
Coeficiente de correlación múltiple	0,9992
Coeficiente de determinación R ²	0,9985
R ² ajustado	0,9746
Error típico	0,2706
Observaciones	45

ANÁLISIS DE VARIANZA

	<i>Grados de libertad</i>	<i>Suma de cuadrados</i>	<i>Promedio de los cuadrados</i>	<i>Valore critico de F</i>	<i>Probabilidad de F</i>
Regresión	3	2025,01205	675,0040178	9216,18971	5,13E-58
Residuos	42	3,07612686	0,073241116		
Total	45	2028,08818			

	<i>Coeficientes</i>	<i>Error típico</i>	<i>Estadístico t</i>	<i>Probabilidad</i>	<i>Inferior 95%</i>	<i>Superior 95%</i>
Intercepción	0	#N/A	#N/A	#N/A	#N/A	#N/A
Ln DAP	1,3363	0,3357	3,9807	0,00026684	0,6588	2,0137
Ln H	0,2196	0,1394	1,5758	0,12256914	-0,0616	0,5009
Ln NA	0,9945	0,0424	23,4743	9,0434E-26	0,9090	1,0800

<i>Inferior 99,0%</i>	<i>Superior 99,0%</i>
#N/A	#N/A
0,4306	2,2420
-0,1564	0,5957
0,8802	1,1088

Según la probabilidad de F el modelo es significativo, con el contraste individual la variable que no es significativo, es la variable Ln de altura Ln H, que debe ser eliminado.

TABLA 12. RESULTADOS DE LA REGRESIÓN MÚLTIPLE

<i>Estadísticas de la regresión</i>	
Coeficiente de correlación múltiple	0,9992
Coeficiente de determinación R ²	0,9984
R ² ajustado	0,9751
Error típico	0,2753
Observaciones	45

ANÁLISIS DE VARIANZA

	<i>Grados de libertad</i>	<i>Suma de cuadrados</i>	<i>Promedio de los cuadrados</i>	<i>Valor crítico de F</i>	<i>Probabilidad de F</i>
Regresión	2	2024,83018	1012,415089	13362,13	1,2850E-59
Residuos	43	3,25800143	0,075767475		
Total	45	2028,08818			

	<i>Coeficientes</i>	<i>Error típico</i>	<i>Estadístico t</i>	<i>Probabilidad</i>	<i>Inferior 95%</i>	<i>Superior 95%</i>
Intercepción	0	#N/A	#N/A	#N/A	#N/A	#N/A
Ln DAP	1,0589	0,2907	3,6422	0,00072225	0,4726	1,6453
Ln NA	1,0175	0,0405	25,1526	2,3815E-27	0,9359	1,0991

<i>Inferior 99,0%</i>	<i>Superior 99,0%</i>
#N/A	#N/A
0,2754	1,8425
0,9085	1,1266

Según \bar{R}^2 ajustado alrededor del 97% de la variación de Ln V en m^3 es explicada por las variables Ln NA y Ln DAP

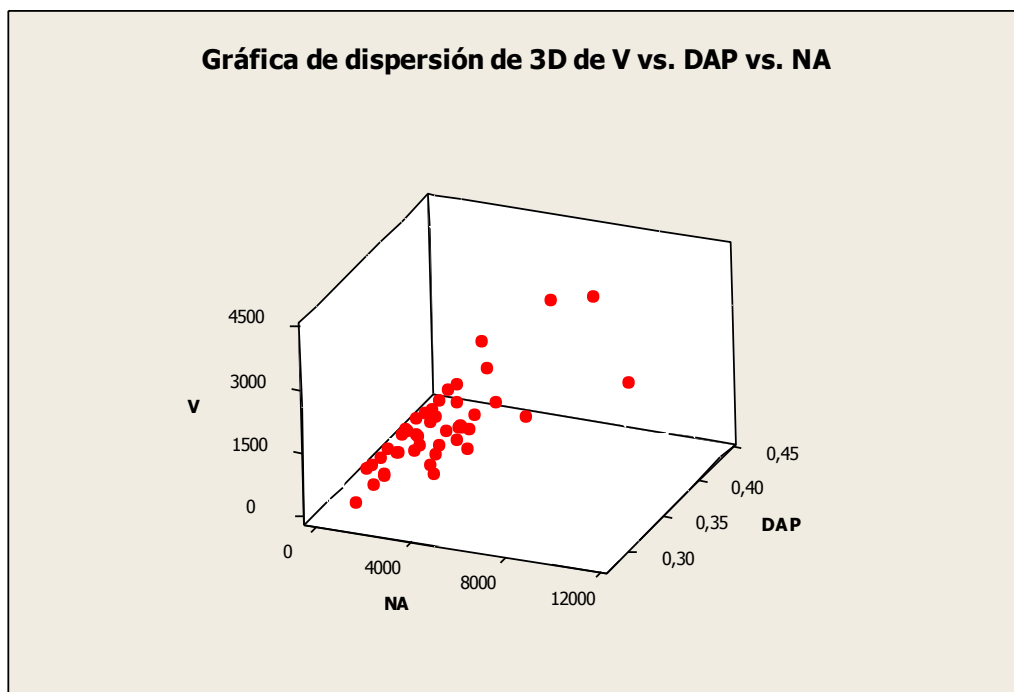
La variable Ln número de árboles es significativo y la variable Ln DAP es significativo, es decir que estas variables aportan al modelo de regresión para el pronóstico.

3.5.3 MODELO BÁSICO DE REGRESIÓN LINEAL

$$\text{Ln } \hat{V} = 1,06 * \text{Ln DAP} + 1,02 \text{ Ln NA}$$

Esta ecuación nos permitirá realizar pronósticos, donde el incremento de 1% en el diámetro a la altura del pecho DAP, en promedio incremento 1,06 % en total de volumen en m^3 en forma similar el incremento de 1% de número de árboles género en promedio un incremento de 1,02 % en el total del volumen en m^3 .

FIGURA 26. DIAGRAMA DE DISPERSIÓN EN 3D



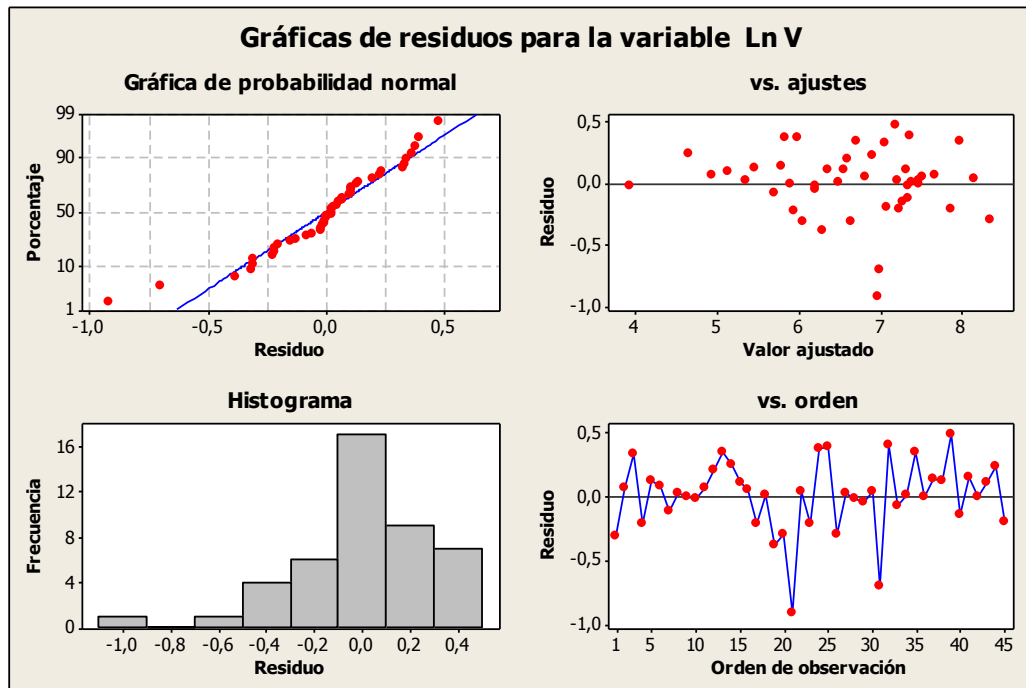
Elaborado por: Tesista

Fuente: MAE-Chimborazo

En la figura 20, la variable volumen es la dependiente, en cambio las variables independientes son: número de árboles NA y diámetro a la altura del pecho DAP, por tanto si tengo más NA y un alto diámetro a la altura del pecho DAP, será mayor volumen en m^3 esto es notable en el diagrama de dispersión en 3D.

3.5.4 ANÁLISIS DE RESIDUOS

FIGURA 27. GRÁFICA DE RESIDUALES



Elaborado por: Tesist

Fuente: MAE-Chimborazo

En la figura 21, los puntos de la gráfica de probabilidad normal, sugiere una buena coherencia entre los datos, en este caso de los residuales, por lo cual se puede concluir que proviene de una distribución normal.

La gráfica de los residuales contra los ajustados se puede ver que la varianza parece no ser constante, para salir de dudas más adelante se comprobara con la prueba de White si existe homocedasticidad.

El histograma tiene una asimetría a la izquierda, tiene sesgo negativo.

El orden de las observaciones presenta aleatoriedad, por lo cual el modelo de regresión es adecuada.

3.5.5 PRUEBA DE WHITE

Para la prueba de White se efectuó la regresión auxiliar

$$\hat{\varepsilon}_i^2 = \alpha_1 + \alpha_2 X_i + v_i$$

Es decir, con el cuadrado de los residuos de la regresión original se hace sobre las variables LN DAP y Ln NA o variables independientes, sobre sus valores al cuadrado y sobre los productos cruzados de las independientes.

TABLA 13. RESULTADOS DE LA REGRESIÓN MÚLTIPLE PARA ERRORES

Estadísticas de la regresión	
Coeficiente de correlación múltiple	0,4736
Coeficiente de determinación R ²	0,2243
R ² ajustado	0,1217
Error típico	0,1512
Observaciones	45

Se obtuvo una R^2 de 0,2243, ahora $n \cdot R^2 = 45(0,2243) = 10,09$, como para distribución ji cuadrado al 5% para 5 df, el valor crítico es 11,07 es mayor que 10,093, se acepta la hipótesis nula, y se concluye con base en la prueba de White que no hay heterocedasticidad.

3.6 PRONÓSTICOS Y EVALUACIÓN CON MEDIDAS DE PRECISIÓN

Se realizó el pronóstico del volumen en m^3 de aprovechamiento de madera, con el modelo de regresión.

Observación: se observe que el modelo tiene la variable en términos de logaritmo al igual que la variable dependiente las mismas que debe utilizar para su interpretación el antilogaritmo (exponencial).

3.6.1 PRONÓSTICO A LARGO PLAZO

TABLA 14. RESULTADO DE PRONÓSTICO A LARGO PLAZO

Mes	Volumen Real	Volumen Pronosticado
1	4071,52	2894,02
2	1765,74	1771,66
3	271,15	238,20
4	1713,17	1530,52
5	2130,18	1314,86
6	1259,15	1454,56
7	378,45	328,88
8	365,39	366,03
9	189,32	170,92
10	3122,34	4505,22
11	1259,87	996,62
12	973,23	1189,57

$$\text{MAD} = 361,73$$

$$\text{MSE} = 346262,26$$

$$\text{MAPE} = 0,1802 \text{ (18,02 \%)}$$

$$\text{MPE} = 0,0426$$

MAD indica que cada pronóstico se desvió en promedio 361,73 m^3 mensuales. El MSE de 346262,26 el MAPE de 18,02% y MPE de 0,0426 indica que la técnica no sobreestima ni subestima el volumen en m^3 mensuales de madera.

3.6.2 PRONÓSTICO MEDIO PLAZO

TABLA 15. RESULTADO DE PRONÓSTICO A MEDIO PLAZO

Mes	Volumen real	Volumen pronosticado
1	378,45	328,88
2	365,39	366,03
3	189,32	170,92
4	3122,34	4505,22
5	1259,87	996,62
6	973,23	1189,57

$$\text{MAD} = 321,85$$

$$\text{MSE} = 338538,98$$

$$\text{MAPE} = 0,1840 \text{ (18,40 \%)}$$

$$\text{MPE} = - 0,0383$$

MAD indica que cada pronóstico se desvió en un promedio de 321,85 m^3 mensuales. El MSE de 338538,98, el MAPE de 18,40 % y por último el MPE es de - 0,0383 indica que la técnica no sobreestima ni subestima el volumen en m^3 mensuales de madera, es el más adecuado en comparación con los pronósticos a largo y corto plazo.

3.6.3 PRONÓSTICO A CORTO PLAZO

TABLA 16. RESULTADO DE PRONÓSTICO A CORTO PLAZO

Mes	Volumen real	Volumen pronosticado
1	3122,34	4505,22
2	1259,87	996,62
3	973,23	1189,57

$$\text{MAD} = 620,82$$

$$\text{MSE} = 676146,07$$

$$\text{MAPE} = 0,2914 \text{ (29,14 \%)}$$

$$\text{MPE} = - 0,1521$$

Tan solo obsérvese el MSE de 6761146,07 y MAPE de 29,14 % los dos valores son altos, por lo que el pronóstico no es el adecuado.

3.7. PRUEBA DE HIPÓTESIS PARA LA VARIABLE VOLUMEN V

1. Establecimiento de hipótesis

$$H_0 : \mu \geq 1689$$

$$H_1 : \mu < 1689$$

2. Selección de un nivel de significancia

Se desea probar para un nivel de significancia del 5% las hipótesis, esto es un $\alpha = 0.05$.

3. Tamaño de la muestra y cálculo de estadístico

Datos

$$n = 54$$

$$\bar{x} = 1689$$

$$s = 1757.33$$

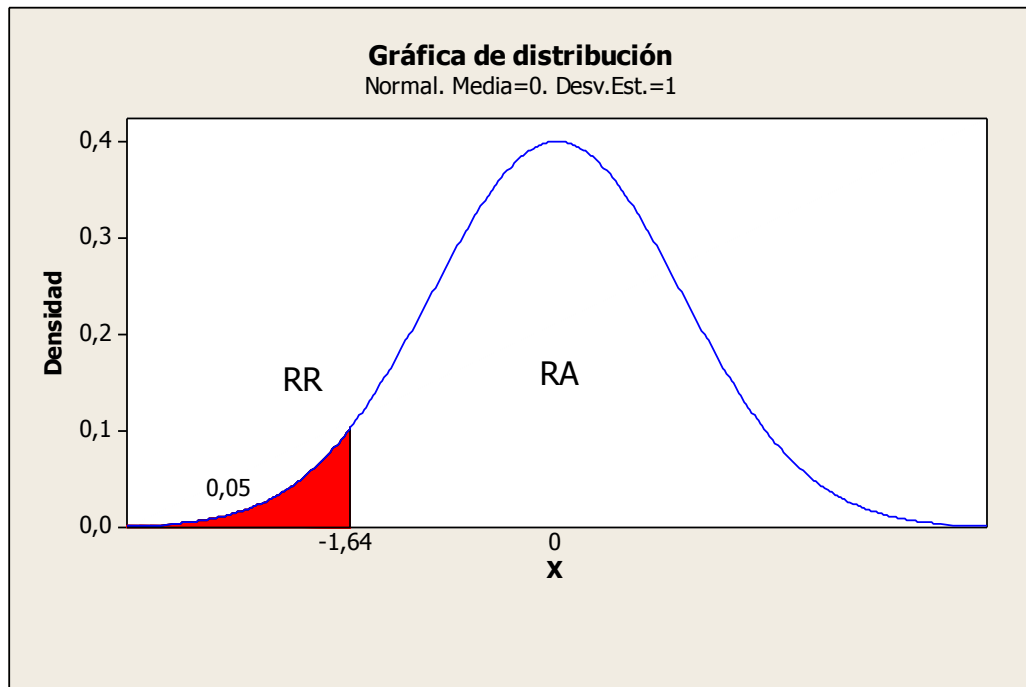
Como $n > 30$, se aplicará el estadístico Z

$$Z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \quad , \quad Z = \frac{1689 - 1689}{1757.33 / \sqrt{54}} = 0$$

4. Especificaciones de las regiones de aceptación y rechazo

Las regiones críticas de aceptación y rechazo se definen por el valor de z , correspondiente a una área de 0.95, tiene un valor de $-1,64$, la misma que se muestra en la figura.

FIGURA 28. GRÁFICA DE LAS REGIONES, ACEPTACIÓN Y RECHAZO



Elaborado por: Tesista
Fuente: MAE-Chimborazo

5. Decisión estadística

Según el valor calculado de $Z = 0$, es mayor que el valor crítico $Z = -1,64$, cae en la zona de aceptación, por lo que se acepta la $H_0: \mu \geq 1689$, se concluye que el

volumen de aprovechamiento de madera de pino será mayor o igual en los próximos meses respecto a la media mensual.

CAPÍTULO V

CONCLUSIONES Y RECOMENDACIONES

4.1 CONCLUSIONES

- El modelo para pronosticar el volumen mensual de la madera de pino está dado por :

$$\ln \hat{V} = 1,06 * \ln \text{DAP} + 1,02 \ln \text{NA}$$

- Para el pronóstico el modelo MBRL adecuado es la logarítmica.
- El modelo de regresión múltiple es altamente significativo (ver tabla 12).
- Las variables de estudio que aportan al modelo son las siguientes: número de árboles (NA) y diámetro a la altura del pecho (DAP).
- La variable altura del árbol (H) no es significativa, la causa es por no aplicar el muestreo estratificado, en las mediciones de las alturas de los árboles.

- Las predicciones son más precisas para pronóstico de medio plazo (para seis meses) en comparación con pronósticos largo y corto plazo (ver tabla 14 – 17).

- El volumen de aprovechamiento de madera de pino será mayor o igual a 1689 m^3 en los próximos meses, según el paso 5 de la prueba de hipótesis para el volumen.

4.2 RECOMENDACIONES

- Implementar una base de datos informáticos en el área forestal del Ministerio del Ambiente Chimborazo.
- Aplicar el método de muestreo estratificado para la medición de la altura del árbol, es decir, por lo menos separar: alto y bajo y calcular por separado el volumen para cada uno de estos dos grupos luego sumar para el volumen total, no promediar.
- Aplicar el hipsómetro para la medición de altura del árbol.
- Promover el uso de modelos de pronósticos para la toma de decisiones, en las diferentes actividades en el MAE- Chimborazo.
- Fomentar la creación de un departamento estadístico en las direcciones Provinciales del Ministerio del Ambiente.

RESUMEN

En la presente investigación se realizó un modelo de pronóstico para el volumen de aprovechamiento de la madera de pino en el Ministerio del Ambiente Chimborazo.

Se aplicó el método analítico, para el análisis de los datos se utilizó el software estadístico Minitab y las herramientas estadísticas de la hoja electrónica Excel. Para la elaboración del modelo primero se especificó las variables: volumen, diámetro a la altura del pecho, factor de forma, altura y número de árboles; con base en los datos históricos mensuales del área forestal del Ministerio del Ambiente de Chimborazo, luego se realizó un análisis de la matriz de correlación y la regresión múltiple. Se concluyó que las variables independientes significativas son: diámetro a la altura del pecho, número de árboles, finalmente se evaluó al modelo mediante el análisis residual, observando que el mismo es adecuado.

Según la evaluación con las medidas de precisión los pronósticos para el volumen mensual, son precisas para medio plazo (para seis meses) con un volumen promedio 1259.54 m^3 , con el 95 % de confiabilidad.

Se recomienda a los técnicos encargados de la medición de las alturas de los árboles deben emplear el muestreo estratificado, con dicha técnica nos ayuda a mejorar el modelo, también se debe implementar una base de datos de los registros forestales.

ABSTRACT

STATICS MODEL ELABORATION WHICH HELP TO FORECAST THE WHOLE PINEWOOD EXPLOTATION VOLUMEN AT THE “ENVIROMENT MINISTRY CHIMBORAZO”

A forecast model for pinewood explotation volume at the “Environment Ministry Chimborazo” was carried out in the present research.

Analytic method was applied; statistic Minitab software and the statics tools of electronic spreadsheet Excel were used for the data analysis. Variables such as: volume, diameter at the height of the chest, for factor, height and number of trees based on the historical monthly data of the forest area of the institution mentioned above were specified for the elaboration of the first model, later a correlation matrix analysis and multiple regression were carried out.

According to the evaluation with the accuracy measurements, the month volume forecasts are accurate for a medium term (for six months) with an average volume 1259.54 m^3 , with 95% of reliability.

It is concluded that the independent variables that are meaningful are: diameter at the height of chest and number of trees. Finally the model was evaluated through residual analysis and it is observed that it is adequate.

It is recommended that the technicians who are in charge of the height-tree measurement use the stratified sampling to improve the model, in addition, a forest record data basis must be implemented.

BIBLIOGRAFÍA

1. **CONGACHA A., JORGE W.**, Estadística aplicada a la educación actividades de aprendizaje., Editorial académica española., Berlín - Alemania., 2012., Pp. 204 - 2047
2. **DIEBOL F. X.**, Elements of forecasting, 3^a.ed. , Cincinnati., OH: South- Western., 2004., Pp. 11- 120
3. **ESPINOZA EDUARDO R.**, Geometría vectorial en R^3 ., Servicios gráficos J.J., Lima – Perú., 2004.,Pp 1- 3.
4. **GUJARATI DAMODAR N., PORTER DAWN C.**, Econometría, 5^a.ed., McGraw Hill., México D.F. – México., 2010., Pp. 188 – 274., 320 - 400
5. **HANKE JOHN E., WICHERN DEAN W.**, Pronósticos en los Negocios., 8^a.ed. Prentice Hall., México D.F. – México., 2006.,

Pp. 57-100

6. **MARTIN G., LABEAGA J. Y MONCHON F.**, Introducción a la Econometría., Prentice Hall., Madrid - España., 1997.,

Pp. 54 - 137.

7. **PEÑA DANIEL.**, Regresión y diseño de experimentos., Alianza Editorial., Madrid - España. , 2002 . , Pp. 341- 423

8. **PEÑA DANIEL**, Análisis de datos multivariantes, McGraw Hill., Madrid - España ., 2002., Pp. 103- 110

9. **PINDYCK ROBERT S., RUBINFELD DANIEL L.**, Econometría, modelos y pronósticos., 4ª.ed., McGraw Hill., México D.F - México., 2000., Pp. 87- 110.

10. **TORRES JUAN R., MAGAÑA OCTAVIO T.**, Evaluaciones de plantaciones forestales., Limusa., México D.F.- México ., 1995., Pp. 60- 71.

11. **WALPOLE R., MYERS R., MYERS S., Y YE K.** Probabilidad y estadística para ingeniería y ciencias., 8ª.ed.,Printice Hall., México D.F. – México., 2007., Pp. 267- 284
12. **MINISTERIO DEL AMBIENTE ECUADOR.**, Aprovechamiento de recursos forestales en el Ecuador y decomisos e infracciones período 2010.,Quito- Ecuador., Pp.1- 25.
13. **PICOS J., COGOLLUDO M.**, Apuntes de dasometría, Universidad de Vigo., Vigo - España ., Pp. 1- 30
14. **RIVAS DANIEL T.**, Evaluación de los recursos forestales, Universidad Autónoma Chapingo, México D.F. - México 2006., Pp.1- 26

BIBLIOGRAFÍA DE INTERNET

1. ANÁLISIS EXPLORATORIO DE DATOS

<http://www.5campus.com>

2012- 07 -06

2. PARA EVALUACIÓN DE RECURSOS FORESTAL

<http://www.ambiente.gob.ec>

2012- 04 - 09

<http://www.rivasdaniel.com>

2012-10-19

3. MANUAL DE INVENTARIO FORESTAL INTEGRADO PARA UNIDADES DE MANEJO

<http://www.wwfca.org>

20120507

<HTTP://WWW.INAFOR.GOB.NI>

2012-05-07

4. FACTOR DE FORMA

<http://www.anam.gob.pa>

2012- 09- 17

ANEXOS

ANEXO 1. MATRIZ DE DATOS MENSUALES

AÑO	MES	V	DAP	FF	H	NA
2008	1	561,38	0,35	0,50	9,67	2040
	2	960,76	0,32	0,52	7,12	2630
	3	1631,92	0,33	0,51	8,16	3270
	4	2096,20	0,33	0,51	6,64	7289
	5	790,20	0,35	0,50	7,82	1884
	6	2311,44	0,33	0,51	7,88	6036
	7	1372,44	0,34	0,51	8,69	4240
	8	216,09	0,34	0,50	6,26	583
	9	50,67	0,34	0,51	7,55	149
	10	484,09	0,34	0,51	7,25	1362
	11	149,34	0,36	0,50	9,53	371
	12	6217,52	0,34	0,51	8,63	13630
2009	13	896,41	0,29	0,53	10,47	2375
	14	1165,23	0,37	0,49	9,92	2109
	15	6820,63	0,48	0,45	6,37	13588
	16	134,59	0,39	0,49	9,77	262
	17	650,89	0,29	0,53	9,22	1917
	18	1946,42	0,28	0,55	9,51	6077
	19	4546,36	0,29	0,53	8,36	14321
	20	1121,68	0,29	0,54	10,11	4454
	21	1628,10	0,31	0,53	8,29	4849
	22	6106,35	0,31	0,52	9,70	13111
	23	373,43	0,37	0,49	7,62	1376
	24	3119,58	0,34	0,52	7,08	11399
2010	25	423,13	0,33	0,52	7,77	2946
	26	1403,23	0,31	0,52	8,59	4069
	27	306,17	0,44	0,47	9,08	805
	28	497,29	0,39	0,49	8,58	837
	29	589,78	0,36	0,49	9,63	1062
	30	314,37	0,28	0,54	6,77	1454
	31	1808,09	0,33	0,51	7,30	4864
	32	1518,56	0,33	0,51	10,11	4289
	33	2730,61	0,27	0,56	8,41	19594
	34	7320,73	0,43	0,48	8,57	6961
	35	473,32	0,34	0,50	8,60	1371
2011	36	3630,01	0,41	0,48	8,46	7756
	37	309,32	0,20	0,60	6,40	2372
	38	3191,83	0,23	0,56	37,63	9537
	39	538,99	0,34	0,50	9,00	2942

	40	2355,27	0,32	0,52	8,35	4614
	41	274,78	0,42	0,47	8,16	660
	42	658,68	0,30	0,53	7,28	2062
	43	4071,52	0,37	0,49	7,28	7043
	44	1765,74	0,32	0,52	8,20	5154
	45	271,15	0,38	0,49	10,24	593
	46	1713,17	0,40	0,48	8,50	3512
	47	2130,18	0,41	0,48	10,22	2948
2012	48	1259,15	0,34	0,51	9,26	3923
	49	378,45	0,40	0,50	11,00	772
	50	365,39	0,43	0,46	11,25	791
	51	189,32	0,33	0,50	11,19	489
	52	3122,34	0,27	0,54	11,48	14968
	53	1259,87	0,34	0,51	12,11	2692
	54	973,23	0,35	0,51	12,60	3140
